MASTER OF SCIENCE BY RESEARCH

The Efficacy of Group and Individual Lifestyle Interventions for Overweight and Obese Adults
a Systematic Review and Meta-Analysis of Randomised Controlled Trials

Abbott, Sally

Award date:
2018

Awarding institution:
Coventry University

The Efficacy of Group and Individual Lifestyle Interventions for Overweight and Obese Adults: a Systematic Review and Meta-Analysis of Randomised Controlled Trials

By

Sally Abbott

May 2018

The Efficacy of Group and Individual Lifestyle Interventions for Overweight and Obese Adults: a Systematic Review and Meta-Analysis of Randomised Controlled Trials

By
Sally Abbott

May 2018

A thesis submitted in partial fulfilment of the University's requirements for the Master of Research

Certificate of Ethical Approval

Applicant:

Sally Abbott

Project Title:

The Influence of Gender on the Efficacy of Group and Individual Lifestyle Interventions for Overweight and Obese Adults: a Systematic Review and Meta-Analysis of Randomised Controlled Trials

This is to certify that the above named applicant has completed the Coventry University Ethical Approval process and their project has been confirmed and approved as Low Risk

Date of approval:

31 January 2017

Project Reference Number:

P51006

Abstract

Introduction

Multi-component lifestyle interventions are recommended for the weight management of adults who are overweight or obese. It is not clear, however, whether delivery individually or in groups influences the efficacy. The objective of this research was to systematically review the effectiveness of group compared to individual lifestyle interventions for weight management.

Methods

The databases MEDLINE, EMBASE, CINAHL, CENTRAL and ISRCTN were searched for published and on-going randomised controlled trials (RCTs) from inception to February 2017. The reference lists of included studies were also searched. Eligible studies were RCTs comparing group against individual lifestyle interventions for weight loss among adults with a BMI >25kg/m^2. Risk of bias was evaluated using the Cochrane Risk of Bias tool. Heterogeneity was investigated using I^2 statistics and sub-group analysis. Meta-analysis primarily used fixed effects methods and either estimated risk ratios or continuous inverse-variance methods.

Results

Eight RCTs with 2,139 participants were identified. Group interventions were 62% more likely to achieve a 5% weight loss relative to individual interventions (RR 1.62, 95% CI [1.40, 1.86], p= <0.00001). Heterogeneity existed (I^2 35%, p= 0.15) and was explained through sub-group analysis by provider (commercial or non-commercial) (p= 0.004). Relative to individual interventions, commercial groups were 89% more likely (RR 1.89, 95% CI [1.58, 2.26], p= <0.00001) while results were similar with no significant difference for non-commercial groups (RR 1.21, 95% CI [0.95, 1.54], p= 0.11), in achieving a 5% weight loss. Commercial groups were marginally more cost-effective (£8,128 per QALY) than non-commercial groups (£8,439

per QALY). Neither group nor individual interventions were favoured when measuring efficacy by markers of cardiovascular (lipid profile, blood pressure) and diabetes (fasting glucose, fasting insulin, HbA1c) disease risk.

Conclusion

There is a moderate degree of certainty that neither non-commercial groups or individual interventions are favoured regarding the likelihood of achieving a 5% weight loss at 1-year. There is a high degree of certainty that commercial group participants are more likely to attain a 5% weight loss at 1-year, compared to individual intervention participants. Both commercial and non-commercial group interventions are cost-effective. Referral to a commercial group intervention should be prioritised over a non-commercial group or individual intervention; if the option is available.

Acknowledgements

I would like to thank my Director of Studies, Dr Deborah Lycett, for all of her patient guidance and support whilst undertaking the MRes programme. Thank you so much for always being just ‘an email away’.

Thank you to University Hospitals Birmingham NHS Foundation Trust for supporting my secondment and thus allowing me to undertake this programme of study part-time over the past two years. I also express my gratitude to the NIHR for providing me with this opportunity and supporting this fellowship financially. I hope this is just the start of my journey.

“Success is only meaningful – and enjoyable – if it feels like your own”

Contents

List of Figures

List of Tables

List of Abbreviations

BMI	Body Mass Index
BOCF	Baseline Observation Carried Forward
CI	Confidence Interval
DRVs	Dietary Reference Values
EQ-5D	EuroQol Five-Dimension Scale
ICER	Incremental Cost-Effectiveness Ratio
IPAQ	International Physical Activity Questionnaire
NRI	Non-Responders Imputation
PHQ-8	Patient Health Questionnaire
RCT	Randomised Controlled Trial
SD	Standard Deviation
SE	Standard Error
SMD	Standardised Mean Deviation
VAS	Visual Analogue Score
VLCD	Very Low Calorie Diet

1.0 Introduction

1.1 Overweight and Obesity

Overweight and obesity is classified by a Body Mass Index (BMI) of >25kg/m^2 and >30kg/m^2, respectively (World Health Organisation 2014). In England, nearly two-thirds (63%) of adults are classified as being overweight or obese according to BMI. Further, obesity is an escalating worldwide epidemic. By the year 2030, it is estimated that nearly 58% of the world's adult population could be either overweight or obese (Kelly et al. 2008).

Obesity impacts all facets of an individual's life. Substantial epidemiologic evidence suggests that a BMI >25kg/m^2 is a risk factor for mortality and morbidity from a number of co-morbidities; including type 2 diabetes, cardiovascular disease, and several cancers (Prospective Studies Collaboration et al. 2009) and reduces life expectancy by up to 8 years (Grover et al. 2015). Beyond physical health, obesity is associated with a reduced quality of life (Jia and Lubetkin 2005) and has a negative influence on social (Westermann et al. 2015) and mental well-being (National Obesity Observatory 2011).

Excess weight has vast economic consequences on society. In the UK, adults with obesity are less likely to be in employment (NICE 2013). Further, those that are in employment have more health-related absence and a lesser productivity (NICE 2012); than adults with a healthy weight. Consequentially, by combining the cost to healthcare and loss of earnings, overweight and obesity is estimated to cost UK society at least £27 billion each year (McKinsey Global Institute 2014).

1.2 Lifestyle Interventions

Obesity is a "complex web" of societal and biological factors which contributes to an individual having a dietary intake that is greater than their energy expenditure; resulting in excess adiposity (The Government Office for Science 2007). Addressing overweight and obesity

poses a significant challenge due to the complexity and interdependency of influencing factors. It is expert opinion, therefore, that interventions need take a holistic approach to the enablement of lifestyle changes (Gortmaker et al. 2011).

There is strong evidence from several systematic reviews that multi-component lifestyle interventions incorporating diet, physical activity and behaviour change leads to greater weight loss, as opposed to physical activity intervention alone (Johns et al. 2014, Greaves et al. 2011, Kirk et al. 2012). Combined behavioural interventions are effective in inducing a weight loss of 5 to 10% (Johns et al. 2014). For this reason, multi-component lifestyle interventions that include behaviour change strategies to increase physical activity, improve eating behaviour and reduce energy intake are the first-line interventions for adults overweight and obesity (NICE 2014a).

Drop-out rates from lifestyle interventions have been reported to be as high as 77% (Finley et al. 2007). Attrition is a consistent and strong predictor of weight loss (Ali et al. 2012) and is considered a surrogate to behavioural adherence. Lesser adherence to weight management behaviours means that lifestyle induced weight loss has a trajectory to diminish over time and can lead to weight re-gain (Butryn et al. 2011). For this reason, a 5 to 10% weight loss should be maintained for more than 12 months to be considered a clinically meaningful and maintained weight loss (Stevens et al. 2006).

1.3 Social Support and Weight Management

Network phenomena suggest that social influences are a factor in the obesity epidemic. It has been proposed that this phenomenon can be harnessed, and that the provision and promotion of positive social support can be used to intervene (Christakis and Fowler 2007). Thoits (1995) delineates social support as a coping resource when handling “stressors”. Social support functions to provide emotional, informational, problem-solving (Thoits 2011) and companionship (Deci and Ryan 2000) furtherance to respond to such stressors.

Social support is positively correlated with weight maintenance after weight loss (Elfhag and Rossner 2005); and is an integral cognitive behavioural approach for weight management. This is a hierarchical behavioural change taxonomy of CALOR-E, whereby social support provision assists individuals in changing their physical activity and eating behaviours (Michie et al. 2013). In turn, this provision paticipants to re-enact specific food and activity behaviours, as outlined within OxFAB taxonomy (Hartmann-Boyce et al. 2016). These include 'buddying support', 'motivational support' and 'imitation modelling' behaviours.

1.3.1 Group Interventions

In relation to weight management, a participant's peers are those who are, alongside, also actively engaged with the weight management programme. Reports suggest that social support, empathy, role modelling, accountability and problem solving are offered in group settings by peers and are important factors for lifestyle change and weight loss (Hayaki and Brownell 1996, Latner et al. 2006).

Based on social-cognitive theory (Bandura 1997), group-based lifestyle interventions capitalise upon participant interactions to promote self-efficacy for behaviour change and thus the attainment of weight loss (Palmeira et al. 2007). Provision of a genuine feeling of empathetic understanding normalises the challenges of weight management and validates participants' concerns and feelings. Perceived autonomous support from peers is associated with greater motivation, manifesting in increased physical activity and healthier eating behaviours (Ng et al. 2014).

1.3.2 Interventionist

Interventionists for individual and non-commercial group weight management programmes tend to be healthcare professionals, who draw from their collective professional experience to help participants and rely on the provision of informational support. Autonomous support from

healthcare professionals has been shown to predict self-regulation and subsequent sustainment of long-term weight loss (Williams et al. 1996).

Healthcare professionals have, until more recently, been the predominant care giver for weight management interventions. However, participants have expressed concern over a perceived lack of sensitivity and understanding from healthcare professionals (Phelan et al. 2015). It has also been highlighted that healthcare professionals have weaker social ties with service users, in comparison with social ties with peers or mentors (Verheijden et al. 2005). It has therefore been put forward that mentor support from lay interventionists may be more effective for weight management (Verheijden et al. 2005).

Commercial weight loss programmes, such as Weight Watchers and Slimming World, recruit 'successful' alumni to lead local programmes; without the requirement for them to have any formal healthcare education. Such coaches are perceived by service users to be more familiar and more accessible (Leahey and Wing 2013). They build trust with service users by sharing their lived experiences, which envisions a desired self in the service user's own future (Markus and Nurius 1986); motivating them towards this goal. In this sense, lay interventionists can serve as aspirational role models for participants (Thoits 2011).

1.4 Existing Literature

Systematic reviews have investigated the efficacy of lifestyle interventions in relation to a primary outcome of weight loss. While similar research exists, the vast majority of current research is not congruent to my own research question. The foci of others' research have been the comparison of active interventions against minimal intervention controls, or have combined group and individual interventions together (Peirson et al. 2014, Hassan et al. 2016, Galani and Schneider 2007, Franz et al. 2007, Booth et al. 2014).

There is, however, some in-direct evidence from Hartmann-Boyce et al. (2014a, 2014b) on the efficacy of interventional components, such as delivery to groups or to individuals. Based on synthesised evidence from RCTs published up to the year 2012, Hartmann-Boyce and colleagues examined efficacy in relation to weight loss. They found that neither group nor individual interventions were associated with weight loss at 1-year. However, the limitation of in-direct evidence is that the results should be interpreted with caution due to the loss of power from randomisation.

A single systematic review exists that does provide direct evidence for the effectiveness of group and individual interventions. This study, authored by Paul-Ebhohimhen and Avenell (2009), provides evidence from RCT studies that group interventions are more effective than individual interventions, as measured by weight loss alone, over 1-year of follow-up. These findings, although directly pooled, are limited in reliability and generalisability owing to the included studies dating back to the pre-1990s.

Current guidelines, (NICE 2014a, 2014b), have been formed based upon the availability of existing quality literature at the time of publication. They offer a brief recommendation that group interventions should be favoured over individual programmes on the basis that these *may* be more cost-effective. This recommendation is based on their virtual economic modelling and was not formed from direct clinical evidence. NICE (2014a) acknowledge that further research is needed to examine the influence of the components of lifestyle programmes on adherence, effectiveness and cost-effectiveness. In particular, NICE (2014a) makes reference to the need for further research to explore the specific components of interventions, including: the interventionist (healthcare professionals versus lay persons), the delivery (individual- versus group-based) and the utilisation of various behaviour change taxonomies.

1.5 Rationale

Current guidance (NICE 2014a, 2014b) on lifestyle programme delivery for weight management is brief and self-admittedly requires further investigation. Periodic evidence synthesis is fundamental in order to update and inform planning on how best to deliver cost-effective health services and outcomes for overweight and obesity.

In the treatment of overweight and obesity, group interventions that offer social support networks may be the foundation to behaviour change for weight management. However, in the absence of any recent evidence synthesis in this area, it is unclear whether group interventions or individual interventions are most effective.

Thus, in the context of existing research in the field, this research provides updated knowledge on the effectiveness of group compared to individual interventions; as measured by weight loss. Further, this research offers a more holistic analysis of efficacy, extending measurement to include patient-reported outcome measures, adverse events, cost-effectiveness and clinical markers of cardiovascular and diabetes disease risk.

1.6 Aims and Objectives

The aim of this systematic review is to establish the effectiveness of group relative to individual lifestyle interventions for the treatment of overweight and obesity in adulthood.

The objectives are to:

(1) Identify relevant RCT studies for inclusion

(2) Critically appraise the methodological quality of included studies

(3) Synthesise data from included studies

(4) Establish the certainty of the evidence using GRADE

(5) Discuss recommendations for clinical practice

(6) Identify priorities for further research

2.0 Methodology

2.1 Study Identification

2.1.1 Eligibility Criteria

The review question was framed in terms of the population (P), intervention (I), comparator (C), outcomes (O) and study design (S) (NHS Centre for Reviews and Dissemination 2009) and is detailed in Table 1.

The review question then informed the inclusion and exclusion criteria; which are outlined in Table 2. These eligibility criteria were pre-defined and published (Abbott, S. and Bryant 2017) prior to carrying out the literature search. The eligibility criteria were detailed explicitly, in order to minimise researcher bias in the screening stages.

Table 1**: Review Question (PICO)**

Population	Adults (>18 years old) Overweight or Obese (BMI >25kg/m2)
Intervention	Multicomponent weight management programme delivered in groups
Comparator	Multicomponent weight management programme delivered individually
Outcomes	Weight change from baseline to 1-year follow-up
Study Design	Randomised Controlled Trials (RCTs)

2.1.2 Literature Search

Literature searches were made using Medical Subject headings (MeSH) or keywords relevant to the framed question of this review. The search strategy was developed and piloted in

consultation with a medical librarian. This was tested and refined in order to achieve the maximum sensitivity for obtaining relevant studies, using an evidence based approach recommended by the Cochrane Collaboration (Glanville et al. 2006). The terms used and how these were combined are detailed in Appendix 1.

A systematic search was conducted using the search strategy on the 9th February 2017. MEDLINE (1946 to present), EMBASE (1974 to present) and CINAHL (1981 to present) were searched via EBSCO Host and the CENTRAL database was searched via The Cochrane Library. The ISRCTN database was searched to identify any research that may not have been published. To ensure all relevant literature was captured, the reference lists of the included studies and one previous systematic review (Paul-Ebhohimhen and Avenell 2009) were also searched.

Table 2: Inclusion and Exclusion Criteria

	Inclusion	Exclusion
Population	● Inclusion BMI ≥25kg/m2 ● ≥ 18 years old	● Inclusion BMI <25kg/m2 ● <18 years old ● Pregnancy
Intervention	● Group/Group + Individual ● Promoting weight loss ● Multi-component programmes ● Non-specific disease management ● Lifestyle intervention only	● Individual only ● Promoting weight maintenance ● Single component programmes ● Disease-specific management ● Surgical, drugs, meal replacements
Comparator	● Individual contact ● Active intervention	● Any group contact ● Minimal intervention/control
Outcome	● Weight change at 1-year	● Weight change not reported at 1-year
Setting	● Outpatient, primary care, community ● Any country	● Inpatients ● No restriction on country
Design	● RCT	● Any other study design
Literature	● Any publication type ● Any year	● No exclusion on publication type ● No exclusion on publication year
Language	● Any language	● No language restriction

2.1.3 Screening Process

2.1.3.1 Titles and Abstract

Two independent reviewers electronically screened all search results for possible inclusion based on title and abstract, using EPPI-Reviewer 4 software. Each reviewer was blinded to the other's decisions until the point of arbitration. Any discrepancies were reviewed by consensus and overseen by a third review author as arbitrator. Search results proceeded to be screened using a full-text assessment when both reviewers were in agreement that the inclusion criteria were met. If there was insufficient detail in the title and abstract, the article was presumed to be inclusive and progressed to full-text screening for further examination.

2.1.3.2 Full-text

Full-text documents were obtained and uploaded electronically onto the EPPI-Reviewer 4 software. Two independent reviewers screened all articles that had progressed to the full-text screening stage. Screening decisions were recorded electronically and each reviewer was blinded to the other's decisions. Reasons for exclusion were given a priori by each reviewer. Reasons for exclusion were, in order, as follows: 'population', 'intervention', 'comparator', 'outcome', 'follow-up period' and 'study design'.

Any discrepancies during the screening of full-texts were discussed between reviewers to reach a consensus. If a consensus was not agreed, a third independent reviewer was involved in arbitration and made the final decision on inclusion.

If there was insufficient detail reported in the full-text, the authors were contacted to extract more information. If the authors did not respond to communication within four weeks, the study was presumed ineligible and was excluded under 'inconclusive'. Likewise, articles were excluded under 'study ongoing' if authors responded but stated that the data had not yet been analysed.

2.2 Data Collection

2.2.1 Data Items

Data was independently extracted by one reviewer, using full-text copies of the included articles. A second author checked the extracted data for any discrepancies (Higgins and Green 2011) Data was extracted onto an electronic version of the 'Good Practice Data Extraction Form' (Cochrane Effective Practice and Organisation of Care 2017) – as shown in Appendix 2. Descriptive information was extracted about the characteristics of each study (e.g. inclusion criteria, participants, study setting, details of the intervention and comparator etc.) to assess for between-study heterogeneity.

Quantitative outcome data was extracted for the primary and secondary outcomes; as detailed below (Table 3), To ensure transparent outcome reporting, all obtainable outcome measures were detailed in a matrix table according to ORBIT classification (Kirkham et al. 2010). Further data was sought and obtained by contacting an author (Jebb et al. 2011). Both of these steps were taken to minimise selective reporting bias within this review.

Table 3: Primary and Secondary Outcome Measures

Outcome	Data Type	Measure	Time-point
Primary outcome			
Proportional weight loss	Binary	Attainment of 5% weight loss	1-year
Secondary Outcomes			
Proportional weight loss	Binary	Attainment of 5% weight loss	Programme-end
Physical activity level	Continuous	IPAQ	1-year, programme-end
Dietary intake	Continuous	Macronutrient composition	1-year, programme-end
Cardiovascular disease risk	Continuous	Blood pressure, lipid profile	1-year, programme-end
Diabetes risk	Continuous	HbA1c, fasting glucose, fasting insulin	1-year, programme-end
Quality of life	Continuous	EQ-5D	1-year, programme-end
Social support[$]	*[$]data not attainable*		
Mental health	Continuous		1-year, programme-end
Cost-effectiveness	Binary	Cost per QALY gained	1-year
Attrition	Binary	Attendance	1-year, programme-end
Risks	Binary	Related adverse events	1-year

2.2.2 Summary Measures

2.2.2.1 Binary Data

Data was extracted as a standard estimation of the risk ratio (RR) and its 95% confidence interval (CI) using frequency data from binary 2x2 tables. RR is a relative effect measure and is known to be more consistent than absolute measures for meta-analysis (Deeks 2002). The RR is defined as the likelihood of exposure to an event (i.e. achievement of 5% weight loss) *relative* to those without exposure to event (i.e. did not achieve 5% weight loss). It is not possible to calculate RR where there is a zero exposure to an event (i.e. adverse events) (Higgins and Green 2011).

2.2.2.2 Continuous Data

The majority of the continuous data for outcome measurement that was extracted was obtained using the same instrument and scale. In these instances, data was extracted as mean, standard deviation (SD), and the number of participants whom contributed data. Where SD were not presented, these were calculated from standard errors (SEs) using the following formula: $SD = SE \times \sqrt{n}$.

While physical activity was measured using the same instrument (IPAQ), the scale varied between studies and thus did not yield comparable data. In order to utilise this data, data was extracted in the same way but was then calculated to an effect size measure using the standardised mean difference (SMD). The limitations of this method are considered and it is acknowledged that SMD is difficult to interpret; given that it is reported in units of standard deviation (Higgins and Green 2011) rather than familiar units.

2.2.3 Missing Data and Imputations

High drop-out rates are inherent within RCTs of obesity interventions (Elobeid et al. 2009). It is assumed to be because of those participants who lose less weight being more reluctant to be followed up. Imputation methods are often used for missing data; such as 'baseline observation carried forward' (BOCF) or 'non-responder imputation' (NRI). BOCF uses the baseline observation, in place of the missing outcome. In the context of weight loss studies, BOCF requires the assumption that all participants who dropped out of the study returned exactly to their baseline weight (Cresswell and Mander 2014). NRI, on the other hand, attributes participant non-attendance with a lack of efficacy (Dossing et al. 2014).

When handling dichotomous outcomes, such as attaining 5% weight loss, we used NRI in the assumption that non-attendance meant a non-achievement of a 5% weight loss. However, the

same is not possible for continuous outcomes. Further, not all included studies in this review used imputations for their missing data and instead reported data for “completers-only”.

We therefore extracted data with a priori preference to BOCF, because it mitigates the bias associated with drop-out rates by assuming a zero weight loss, but also included extracted completers-only data where this was not available.

2.2.4 Studies with Multiple Treatment Groups

Some of the included studies contributed more than one intervention group to the meta-analyses. To overcome unit-of-analysis error, we formed intervention groups with pair-wise comparisons, as recommended by Cochrane methods (Higgins and Green 2011). For dichotomous outcomes, the sample sizes and number of events were summed across groups. For continuous outcomes, means and standard deviations were combined using the formulae outlined in Appendix 4.

Data was extracted separately for each intervention (i.e. group) while the comparator interventions (individual) were combined, where applicable. Once combined, the comparator group was split equally to avoid double counting the pooled results (Higgins and Green 2011).

2.2.5 Risk of Bias Assessment (Individual Studies)

Risk of bias was assessed in accordance with Cochrane guidelines and used an adapted version of the Cochrane Risk of Bias Tool (Higgins et al. 2011). This was adapted by removing the ‘blinding of participants and personnel’ item. This was removed because it is not feasible to blind interventionists and participants to lifestyle interventions.

Included studies were assessed independently by two reviewers for risk of bias. Any disagreements were resolved by discussion between reviewers and if necessary, involved arbitration by the third reviewer. Both reviewers independently assigned judgement of high,

low or unclear risk for each item for all included studies, along with a free text justification (as shown in Appendix 3).

For the purpose of this systematic review, it was judged that the key domains that could influence the integrity of the results were 'incomplete outcome data', 'random sequence generation' and 'allocation concealment'. Therefore, when assigning an overall risk of bias for each study, all three items must have been judged to be 'low' risk for a study to be judged to have a 'low' risk of bias overall. If any of these three items are judged to have an 'unclear' or 'high' risk of bias; the study was classified as having an 'unclear/high' risk of bias. This overall classification of a study's risk of bias was then utilised as part of GRADE assessment, and for the pre-defined sensitivity analyses.

2.3 Data Synthesis

The data synthesis methodology was pre-defined in the protocol (Abbott and Bryant 2017). This ensured that post-hoc analyses were not performed, therefore avoiding the induction of biases associated with selective reporting of only statistically significant results.

2.3.1 Heterogeneity

A visual test for heterogeneity was used to assess the overlap in confidence intervals for each effect estimate on a forest plot. If the overlap was poor or there were outliers, a test for statistical heterogeneity was performed. Statistical assessment of heterogeneity used the I^2 method alongside the Chi^2 p-value. I^2 provides an estimate of the percentage of inconsistency thought not to be due to chance. Substantial statistical heterogeneity was defined as an I^2 of above ~50%, and/or accompanied by a statistically significant Chi^2 p-value (Higgins et al. 2011).

2.3.2 Approaches

Meta-analysis was undertaken for each outcome where data had been extracted from a minimum of 2 studies (Valentine et al., 2010) using RevMan 5 software.

Fixed-effects meta-analysis, using Mantel-Haenszel methods (Mantel and Haenszel 1959) was used when the observed differences among study results were solely due to chance; i.e. where substantial statistical heterogeneity was not identified, or where heterogeneity was explained by sub-group analysis.

Meanwhile, random-effects meta-analysis, using DerSimonian and Laird methods, (DerSimonian and Laird 1986), was used to incorporate heterogeneity among studies, only when heterogeneity was un-explained by sub-group analysis (Higgins and Green 2011).

A narrative approach was taken instead, if data was obtainable for only one included study or if substantial heterogeneity was found (see 2.3.1). Narrative synthesises explored within and between-study results in line with the guidance from the Centre for Reviews and Dissemination (2009).

2.3.3 Meta-biases

Testing for publication bias using funnel plot asymmetry was not appropriate given that there were only 8 studies included in the review. This meant that the test power would be too low to distinguish chance from genuine asymmetry (Sterne et al. 2011). Publication bias was instead explored descriptively and according to classification of sample size (Easterbrook et al. 1991). For this exploration, the pre-specified primary outcome for each included study was used to assess for bias in publication according to the significance ($p = <0.05$) of results.

Larger studies in this review were conducted with greater methodological rigor. Further, only one small study was included in this review (Long et al. 1983). This singular study contributed

a very low weighting to the fixed-effects meta-analysis and only had usable data for one secondary outcome alone (see 3.5.10). Based on this, small-study effects are unlikely to have influenced the results (Higgins and Green 2011).

2.3.4 Confidence in Cumulative Evidence

Six outcomes were identified as being the most important outcome measures for patient-care and decision making, and are outlined in Table 4.

The GRADE approach (Schünemann et al. 2008) was used to interpret findings on the basis of evidence certainty. Each outcome was assessed against the following domains: risk of bias, inconsistency, imprecision, indirectness and publication bias. GRADEPRO was used to create a summary of findings table, which also communicated the magnitude and sum of the absolute effect of the interventions included in the review. We presented these absolute effects using baseline risk, risk difference and the number needed-to-treat (NNT) for a beneficial outcome. Communicating absolute effects is important in the interpretation of meta-analysis findings as they are more intuitive (Newcombe and Bender 2014).

Table 4: Rationale for Inclusion of GRADE Outcomes

Outcome	Rationale
Achievement of 5% weight loss	Objective clinical measure: internationally accepted target which demonstrates health improvements (Jensen et al. 2014, NICE 2014a)
Systolic blood pressure	Objective clinical measure: continuous marker of cardiovascular disease risk (World Health Organisation 2002, NICE 2011)
Total:HDL cholesterol	Objective clinical measure: measure of both total and HDL cholesterol is recommended by (NICE 2016) as the best measure of CVD risk
HbA1c	Objective clinical measure: reflects long-term glycaemic exposure (World Health Organisation 2011)
Quality of Life	Patient reported outcome measure: emotional and psychological assessment of well-being (Janse et al. 2004)
Cost-effectiveness	Health economics measure: assessment of whether the intervention provides significant benefit at an acceptable cost (Appleby et al. 2007)

2.4 Pre-Specified Additional Analyses

Sensitivity analyses were performed by excluding studies which were judged to have an 'unclear' or 'high' risk of bias, see 2.3.5. These were undertaken to determine whether the results were affected by study quality. Where results were not altered by sensitivity analysis, the results of the studies, regardless of risk of bias, can still be regarded with a degree of certainty. If, conversely, results were affected by the sensitivity analysis; the original analysis of all included studies will be interpreted with caution. Instead, results obtained from the analysis of only 'low' risk of bias studies will have more credibility.

The primary outcome had a near statistically substantial heterogeneity. As was pre-specified in the protocol, sub-group analysis was therefore performed to explain this. Study characteristics were inspected to identify potential methodological or clinical heterogeneity. Sub-group analysis by the intervention provider (categorised as being commercial or non-commercial) explained heterogeneity entirely for the primary outcome. Thus, this sub-group analysis was then also performed on the meta-analyses of the secondary outcomes.

3.0 Results

3.1 Study Selection

The study selection process is detailed in Figure 1. A total of 8 studies were identified for inclusion in the review. These were generated from a search of the databases MEDLINE, CINAHL, EMBASE, CENTRAL and ISRCTN, which retrieved a total of 5,856 citations. Importing the reference lists of the 8 included studies and the relevant systematic review by Paul-Ebhohimhen and Avenell (2009) provided an additional 485 citations. After adjusting for duplicates, the citations identified in the literature search totalled 5,673. Of those, 5,467 citations were discarded after screening their titles and abstracts as they did not meet the inclusion criteria. The full texts of the remaining 206 citations were then examined in further detail.

After examination of the full-text, 198 citations were excluded because they did not meet the inclusion criteria. The reasons for excluding studies are given in detail in Figure 1. An inappropriate comparator was the main reason for exclusion of full-text articles (n= 104, 50.4%), usually because the comparator was a minimal intervention or control, or involved group-delivery. Other reasons for exclusion were study design (before-after studies or secondary analyses), population (entry BMI unspecified or including participants with a BMI <25kg/m^2) and intervention (involving pharmacological, VLCD or surgical interventions; or were not multi-component).

Figure 1: PRISMA Flowchart

Some materials have been removed due to 3rd party copyright. The unabridged version can be viewed in Lancester Library - Coventry University.

Reproduced from Moher et al. (2009)

3.2 Study Characteristics

3.2.1 Participants

A combined total of 2,139 overweight or obese adult participants were included in this review overall (refer to Table 2). The number of participants in each study ranged from 36 to 779. All the studies included were conducted in westernised countries; of which 50% were conducted within UK populations. Seven of the 8 studies included both male and female participants, while one study by Long et al. (1983) included only females. The majority (75%) of studies reflected a mean age of between 45 to 50 years old.

Inclusion BMI criteria across studies varied considerably. Five studies included participants who were overweight or obese (>25kg/m^2 to >28kg/m^2); while the remaining studies included only participants who were classified as obese (>30kg/m^2 to >40kg/m^2). Despite the inclusion criteria being accessible to participants who were classified as overweight, the mean BMI for all studies was >30kg/m^2, and thus obese according to BMI classification.

3.2.2 Contact Time

The duration of study follow-up was 12 or 24 months; but the majority (75%) ran for 12 months. In all of the group interventions, the intensive initial phase consisted of weekly sessions, whereas only half of individual interventions were delivered weekly. Where stated, the total contact time for a participant in the group intervention was greater (12 – 55.3 hours) than for a participant in the individual intervention (2.5 – 11 hours).

3.2.3 Setting

Out of the 12 group interventions, most (n= 5) were performed in the commercial slimming club setting and the least were performed in primary care (n= 1) and in the community setting (n= 1). Out of the 9 comparator arms, most were conducted in the primary care (n= 3) and out-patient (n= 3) settings. All the group interventions were delivered in-person; whereas one

of the individual interventions was delivered remotely via telephone and email (Appel et al. 2011).

3.2.4 Interventionist

Among the group interventions, more non-healthcare professionals (5 commercial Slimming Club Leaders, 1 Food Advisor, 1 Weight Loss Coach) than healthcare professionals (1 Psychologist, 1 Specialist Nurse, 2 Dietitians, 1 Joint Dietitian/Psychologist) delivered the intervention. Notably, 1 of these interventions was provided through Slimming World (Jolly et al. 2011) and 3 interventions were provided by Weight Watchers (Jolly et al. 2011, Heshka et al. 2003, Jebb et al. 2011). These two commercial programmes are exclusively led by prior members who have been “successful” with weight loss.

For the individual interventions, nearly all interventionists were healthcare professionals (3 Primary Care Nurses/GPs, 1 Pharmacist, 3 Dietitians, 1 Dietitian & Endocrinologist) and one interventionist was a non-healthcare professional (Weight Loss Coach).

3.2.5 Behavioural Taxonomies

The underpinning behavioural components were described in sufficient detail to assign clustered behavioural taxonomies (Michie et al. 2013) to 83% (n= 10) of the intervention arms and to 67% (n= 6) of comparator arms. The group interventions utilised between 5 and 13 different taxonomies and the individual interventions utilised between 4 and 8 different taxonomies. Across both group and individual interventions, the most frequently utilised taxonomies were Goals and planning (100%), Feedback and monitoring (100%) and Comparison of outcomes (100%). In addition, 100% of group interventions utilised Social support taxonomies, in contrast with 50% of individual interventions.

Table 5: Study Characteristics

Author	Year	Country	N	Inclusion BMI	Intervention		Comparator		BMI (kg/m^2)	Age (years)	% male
					Interventionist	Details	Interventionist	Details			
Appel	2011	USA	277	30-50	Weight Loss Coach	○ Primary care setting ○ Delivered in-person + remote telephone/email ○ 24 months duration: weekly (0-3 months), monthly (4-6months), bi-monthly (7-24months) ○ 55.3hrs contact time ○ Behavioural taxonomies [1,3,8,9,10,15]	Weight Loss Coach	○ Remote setting ○ Delivered remotely telephone/email ○ 24 months duration: weekly (0-3 months), monthly (4-6months), bi-monthly (7-24months) ○ 11hrs contact time ○ Behavioural taxonomies [1,3,8,9,10,15]	36.4	55	36
Ash	2006	Australia	128	>27	Dietitian	○ Hospital outservice user setting ○ Delivered in-person ○ 12 months duration: weekly (0-1.5 months), monthly (2-6 months), once (12 months) ○ 14hrs contact time ○ Behavioural taxonomies [1,3,6,8,9,10,13,15]	Dietitian	○ Hospital outservice user setting ○ Delivered in-person ○ 7.5 months duration: weekly (0-2 months), monthly (3-6 months), once (12 months) ○ Contact time not stated ○ Behavioural taxonomies not described	34.0	49	29
Heshka	2003	USA	423	27-40	Commercial Group Leader (Weight Watchers)	○ 'Slimming club' setting ○ Delivered in-person ○ 24 months duration: weekly (0-3 months), bi-annually (4-24 months) ○ Contact time not stated ○ Behavioural taxonomies [1,2,3,8,9,11,12,13,14,15]	Dietitian	○ Setting not stated ○ Delivered in-person ○ 24 months duration: twice (0-3 months), bi-annually (4-24 months) ○ 2.5hrs contact time ○ Behavioural taxonomies [1,8,11,15]	33.7	45	16
Jebb	2011	UK, Australia, Germany	772	30-35	Commercial Group Leader (Weight Watchers)	○ 'Slimming club' setting ○ Delivered in-person ○ 12 months duration: weekly (0-12 months) ○ Contact time not stated ○ Behavioural taxonomies [1,2,3,8,9,11,12,13,14,15]	Nurse or GP	○ Primary care setting ○ Delivered in-person ○ 12 months duration: monthly (0-12 months) ○ Contact time not stated ○ Behavioural taxonomies [1,8,11,15]	31.4	48	13
Jolly	2011	UK	540	>25	Commercial Group Leader (Weight Watchers)	○ 'Slimming club' setting ○ Delivered in-person ○ 12 months duration: weekly (0-3 months), once (12 months) ○ 12hrs contact time ○ Behavioural taxonomies [1,2,3,8,9,11,12,13,14,15]	Pharmacist	○ Pharmacy setting ○ Delivered in-person ○ 12 months duration: weekly (0-3 months), once (12 months) ○ 3.75hrs contact time ○ Behavioural taxonomies [1,2,3,8,9,11,14,15]	33.6	50	32
					Commercial Group Leader (Slimming World)	○ 'Slimming club' setting ○ Delivered in-person ○ 12 months duration: weekly (0-3 months), once (12 months) ○ 18hrs contact time ○ Behavioural taxonomies [1,2,3,5,6,8,9,11,14,15]	Primary Care Nurse	○ Primary care setting ○ Delivered in-person ○ 12 months duration: weekly (0-3 months), once (12 months) ○ 3.75hrs contact time ○ Behavioural taxonomies [1,2,3,8,9,11,14,15]			

Table 6: Study Characteristics (continued)

Author	Year	Country	N	Inclusion BMI	Intervention		Comparator		BMI (kg/m^2)	Age (years)	% male
					Interventionist	Details	Interventionist	Details			
Jolly					Commercial Group Leader (Rosemary Conley)	○ Slimming club' setting ○ Delivered in-person ○ 12 months duration: weekly (0-3 months), once (12 months) ○ 18hrs contact time Behavioural taxonomies [1,2,8,9,10.11.12.13.14,15]					
					Food Advisor	○ Community setting ○ Delivered in-person ○ 12 months duration: weekly (0-1.5 months), once (12 months) ○ 12hrs contact time ○ Behavioural taxonomies [1,8,9,10,11,14,15]					
Long	1983	UK	36	>25	Dietitian	○ Hospital outservice user setting ○ Delivered in-person ○ 12 months duration: weekly (0-4 months), quarterly (6-12 months) ○ 14hrs contact time ○ Behavioural taxonomies not described	Dietitian	○ Hospital outservice user setting ○ Delivered in-person ○ 12 months duration: weekly (0-4 months), quarterly (6-12 months) ○ 5.5hrs contact time ○ Behavioural taxonomies not described	33.5	37	0
					Dietitian & Psychologist	○ Hospital outservice user setting ○ Delivered in-person ○ 12 months duration: weekly (0-4 months), quarterly (6-12 months) ○ 20hrs contact time ○ Behavioural taxonomies not described					
McRobbie	2016	UK	330	28-45	Psychologist	○ Primary care setting ○ Delivered in-person ○ 12 months duration: weekly (0-2 months), monthly (3-12 months) ○ 18hrs contact time ○ Behavioural taxonomies [1,2,3,4,5,6, 8,9,10,11,12,14,15]	Primary Care Nurse	○ Primary care setting ○ Delivered in-person ○ 12 months duration: fortnightly (0-2 months), bi-annually (6-12 months) ○ 3hrs contact time ○ Behavioural taxonomies [1,8,11,15]	35.4	46	29
Tur	2013	Spain	106	>40	Specialist Nurse	○ Hospital outservice user setting ○ Delivered in-person ○ 12 months duration: weekly (0-3 months), bi-weekly (4-12 months) ○ 46.5hrs contact time ○ Behavioural taxonomies [1,3,4,8,15]	Dietitian & Endocrinologist	○ Hospital outservice user setting ○ Delivered in person ○ 12 months duration: pragmatic (0-12 months) ○ Contact time not stated ○ Behavioural taxonomies not described	46.2	48	33

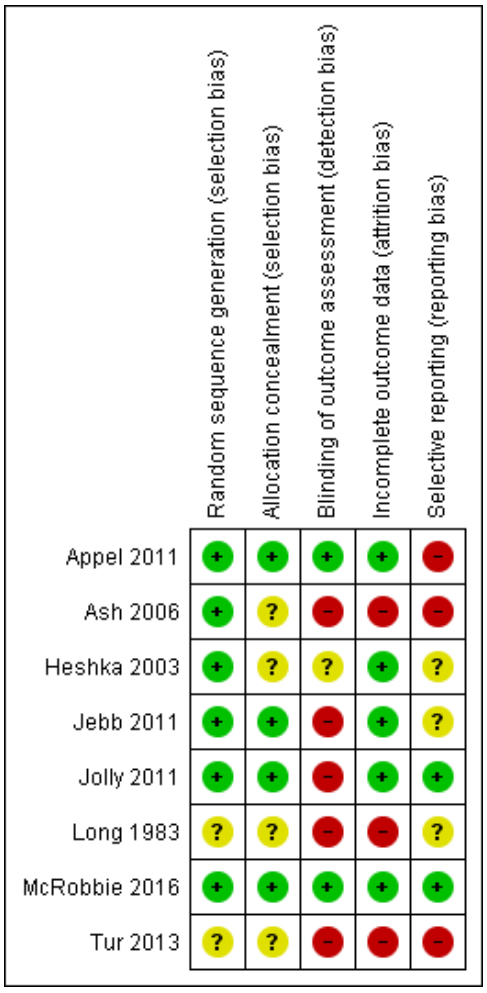
Behavioural taxonomies: [1]Goals and planning, [2]Reward and threat, [3]Regulation, [4]Antecedents, [5]Identity, [6]Self-belief, [7]Covert learning, [8]Feedback and monitoring, [9]Social support, [10]Shaping knowledge, [11]Natural consequences, [12]Comparison of behaviour, [13]Associations, [14]Repetition and substitution, [15]Comparison of outcomes, [16]Scheduled consequence

3.3 Risk of Bias within Studies

The risk of bias across the domains within each study shows variation in the quality of the included studies (Figure 2). McRobbie et al. (2016) was at 'low' risk of bias across all domains. Conversely, Long et al. (1983) and Tur et al. (2013) had no domains considered to be at 'low' risk of bias.

An assessment of the overall risk of bias of individual studies classified four studies with a 'low' risk of bias (McRobbie et al. 2016, Jebb et al. 2011, Jolly et al. 2011, Appel et al. 2011), one study with an 'unclear' risk of bias (Heshka et al. 2003) and three studies with a 'high' risk of bias (Long et al. 1983, Ash et al. 2006, Tur et al. 2013).

Figure 2: Risk of Bias of Individual Studies



3.4 Outcomes reported in Individual Studies

The outcome matrix, as shown in Figure 3, illustrates outcome reporting for individual studies at the 1-year follow-up programme-end time points.

3.4.1 1-year follow-up

Weight change was stated as the primary outcome of measure across all studies. Despite weight change being reported in kilograms in all eight studies, weight change as a proportion (%) was reported in only 5 of these studies (Appel et al. 2011, Heshka et al. 2003, Jebb et al. 2011, Jolly et al. 2011, McRobbie et al. 2016). The most consistently reported outcome was attrition, whereby data was provided for all eight studies. Other than the outcomes of weight change and attrition, the outcomes reported by individual studies varied considerably.

Adverse events were reported by half (n= 4) of studies (Appel et al. 2011, Heshka et al. 2003, Jebb et al. 2011, McRobbie et al. 2016). Some outcomes were reported by only one study each including; dietary intake (Jebb et al. 2011), and quality of life and mental health (Appel et al. 2011). Data on quality of life was collected by a further two studies (McRobbie et al. 2016 and Jebb et al. 2011), but was presented only as part of cost-effectiveness analyses and thus extraction was not possible. Data was not reported by any study for the outcome of social support.

For cardiovascular outcomes, biochemical markers of lipid profiles were recorded for half (n= 4) of studies (Appel et al. 2011, Heshka et al. 2003, Jebb et al. 2011, Tur et al. 2013); although the markers measured varied by individual study (see Figure 3). Meanwhile, systolic and diastolic blood pressure readings were reported by 63% (n= 5) of studies (Appel et al. 2011, Heshka et al. 2003, Jebb et al. 2011, McRobbie et al. 2016, Tur et al. 2013).

Figure 3: Outcome Matrix for Individual Studies

	Weight change (kg)	Weight change (%)	5% weight loss attained	Quality of life	Dietary intake	Physical activity	Social support	Mental health	Blood Pressure	Lipids: Total-C	Lipids: HDL-C	Lipids: LDL-C	Lipids: Total: HDL-C	Lipids: Triglycerides	Fasting Glucose	Fasting Insulin	HbA1c	Attrition	Related Adverse Events	Cost-effectiveness
1-year follow-up																				
Appel (2011)	✔	✔	✔	✔	✘	✘	✘	✔	✔	✔	✔	✔	✘	✔	✔	✘	✘	✔	✔	✘
Heshka (2003)	✔	o[1]	✔	✘	✘	✘	✘	✘	✔	✔	✔	✘	✔	✔	✔	✔	✘	✔	✔	✘
Jebb (2011)	✔	o[1]	✔	o[2]	✔	✔	✘	✘	✔	✔	✘	✘	✔	✔	✔	✔	✔	✔	✔	✔
Jolly (2011)	✔	o[1]	✔	✘	✘	✔	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✘	✘
McRobbie (2016)	✔	o[1]	✔	o[2]	✘	✔	✘	✘	✔	✘	✘	✘	✘	✘	✘	✘	✘	✔	✔	✔
Ash (2006)	✔	o[1]	o[1]	✘	✘	✔[3]	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✘	✘
Long (1983)	o[4]	o[1]	o[1]	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	✘	✘
Tur (2013)	✔	o[1]	o[1]	✘	✘	✘	✘	✘	✔	✔	✔	✔	✘	✔	✔	✘	✔	✔	✘	✘
Programme-end																				
Appel (2011)	✔	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Heshka (2003)	✔	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	N/A	N/A
Jebb (2011)	✔	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Jolly (2011)	✔	o[1]	✔	✘	✘	✔	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	N/A	N/A
McRobbie (2016)	✔	o[1]	✔	✘	✘	✔	✘	✘	✔	✘	✘	✘	✘	✘	✘	✘	✘	✔	N/A	N/A
Ash (2006)	✔	✘	✘	✘	✘	✔[3]	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	N/A	N/A
Long (1983)	o[5]	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✘	✔	N/A	N/A
Tur (2013)	✔	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

✔indicates fully reported results; o indicates partially reported; ✘indicates not reported

[1] weight reported in kg, [2]reported only within cost-effectiveness analysis, [3]non-validated arbitrary scoring, [4]unable to calculate SD, [5]reported as median

At least one marker of diabetes risk was measured for half (n= 4) of studies; of which all reported on fasting glucose (Appel et al. 2011, Heshka et al. 2003, Jebb et al. 2011, Tur et al. 2013). Two of these studies each reported on fasting insulin (Heshka et al. 2003, Jebb et al. 2011) and HbA1c (Jebb et al. 2011, Tur et al. 2013).

Physical activity was measured by four studies (Jebb et al. 2011, Jolly et al. 2011, McRobbie et al. 2016, Ash et al. 2006); all of which used the IPAQ questionnaire for measurement. However, Ash et al. (2006) used a non-validated arbitrary measurement of being 'sufficiently active' which was subjective to the authors and thus was not extracted. The three remaining studies reported through validated methods of scoring and the reporting units varied across mins/week, kcals/week, sitting time/week and MET-mins/week.

3.4.2 Programme-end

For three studies, Jebb et al. (2011), Tur et al. (2013) and Appel et al. (2011), programme-end outcome measurements were not applicable because these interventions were administered over the entire 1-year period. The five remaining studies (Heshka et al. 2003, Jolly et al. 2011, Long et al. 1983, Ash et al. 2006) all **reported** on attrition at programme-end. Of these five studies, only McRobbie et al. (2016) and Jolly et al. (2011) reported further on clinical outcomes. Both studies reported on proportion of participants achieving greater than 5% weight loss and physical activity level change, while McRobbie et al. (2016) in addition reported on blood pressure changes.

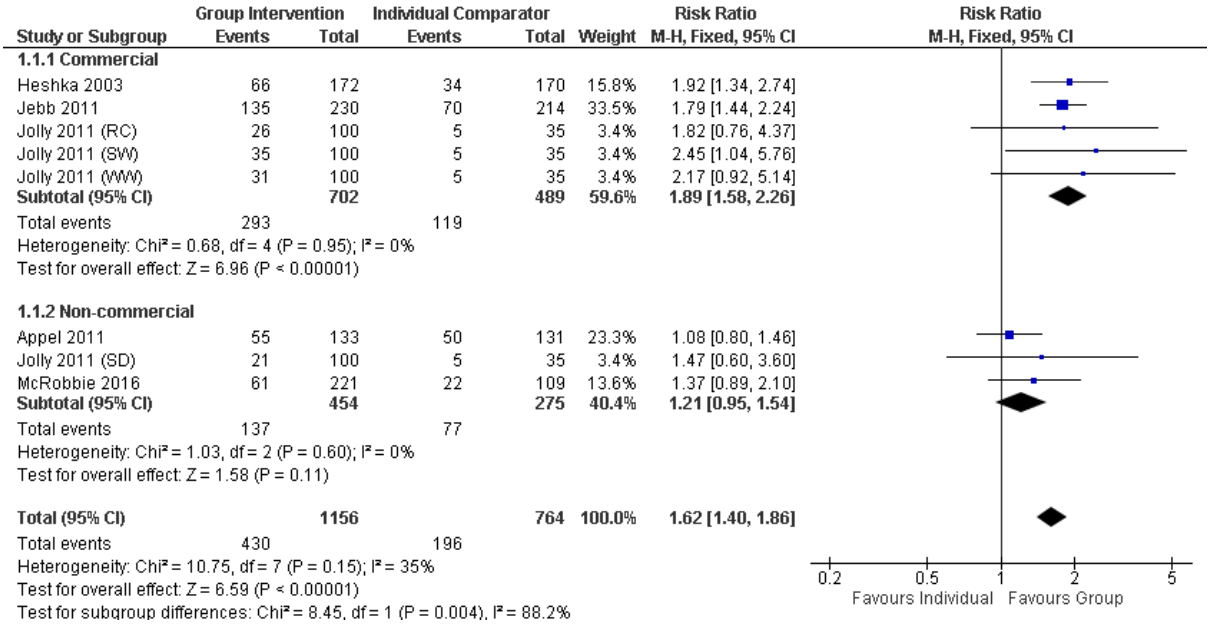
3.5 Meta-Analysis of Results

The findings of this review are based upon 12 group interventions across 8 RCTs. Five group interventions were commercial (888 participants), 7 were non-commercial (605 participants) and these were compared with 9 corresponding individual interventions (1,119 participants).

3.5.1 Weight Loss

The likelihood of achieving 5% weight loss was analysed in 5 studies at the 1-year follow-up time-point (Figure 4). Results showed that participants of group interventions were 62% more likely to attain 5% weight loss at 1-year (RR 1.62, 95% CI [1.40, 1.86], p= <0.00001) relative to individual interventions. Some heterogeneity remained and was dealt with by sub-group analysis. There was a significant difference between the non-commercial and commercial group analyses (p= 0.004) (Figure 4).

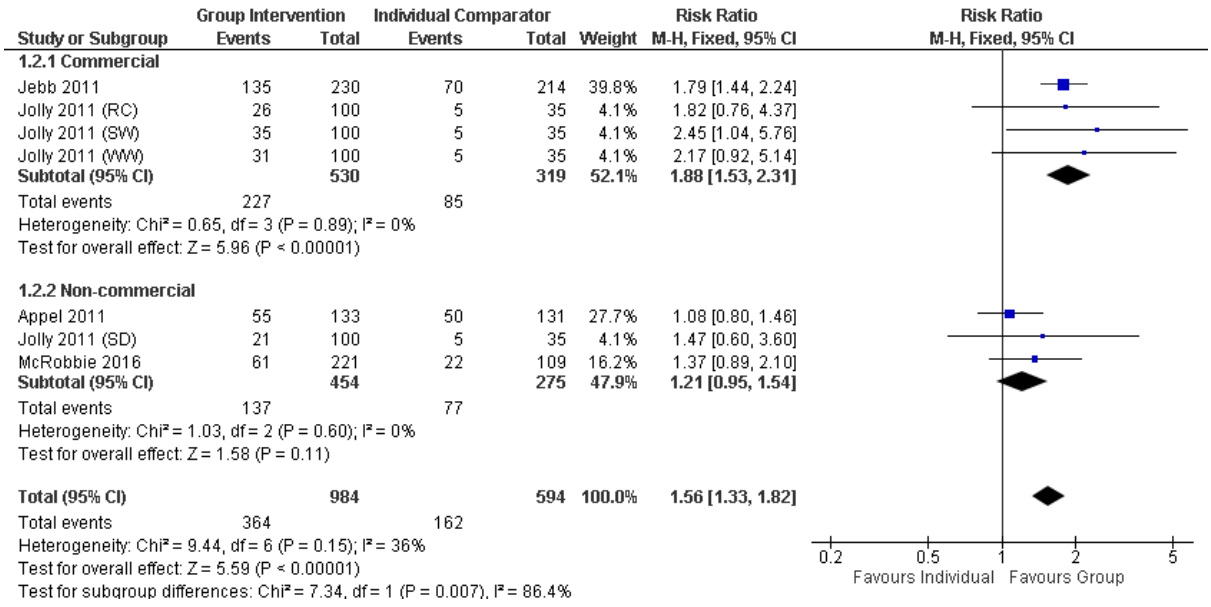
Figure 4: Achievement of 5% Weight Loss at 1-year (All Studies)



Attendees of commercial group interventions were 89% more likely to achieve 5% weight loss (RR 1.89, 95% CI [1.58, 2.26], p= <0.00001) relative to attendees of individual intervention (Figure 4). Meanwhile, non-commercial group attendance meant a 21% increased likelihood of achieving 5% weight loss (RR 1.21, 95% CI [0.95, 1.54]) relative to attendance at an individual intervention, but this was not significant (p= 0.11) (Figure 4). When considering only those studies with low risks of bias, results were similar (Figure 5). These findings support the conclusion that commercial group intervention is superior to individual intervention, in relation

to outcomes of 5% weight loss. Meanwhile, neither non-commercial group interventions nor individual interventions were favoured.

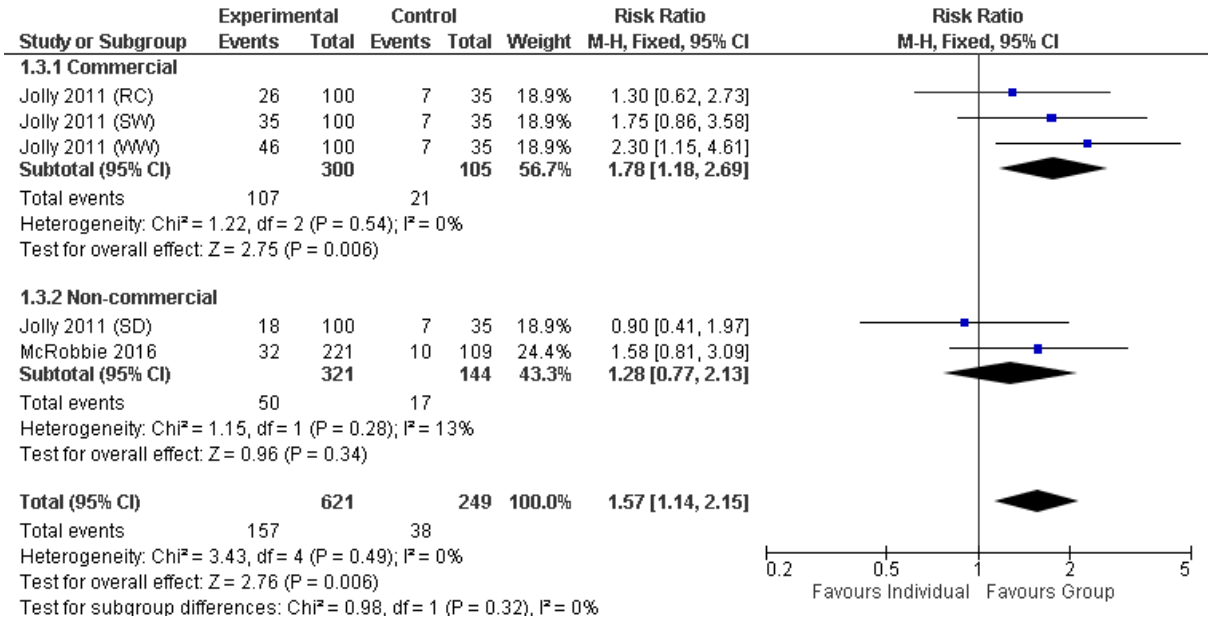
Figure 5: Achievement of 5% Weight Loss at 1-year (Low Risk of Bias Studies)



Two studies measured likelihood of achieving more than 5% weight loss at the programme-end time-point (Figure 6). Participants of group interventions were 57% more likely to achieve 5% weight loss, relative to individual interventions (RR 1.57, 95% CI [1.14, 2.15], p= 0.006). Although these results were statistically homogenous, pre-specified sub-group analysis was performed by provider. Participants of commercial group interventions were 78% more likely to attain a 5% weight loss relative to individual interventions (RR 1.78, 95% CI [1.18, 2.69], p= 0.006). Meanwhile, neither non-commercial nor individual intervention was favoured (RR 1.28, 95% CI [0.77, 2.13], p= 0.34). A sensitivity analysis was not performed as the included studies were judged as a low risk of bias.

Therefore, based on these findings from exclusively low risk of bias studies, commercial groups are favourable in attaining 5% weight loss outcomes relative to individual interventions. However, neither non-commercial groups nor individual interventions were favoured.

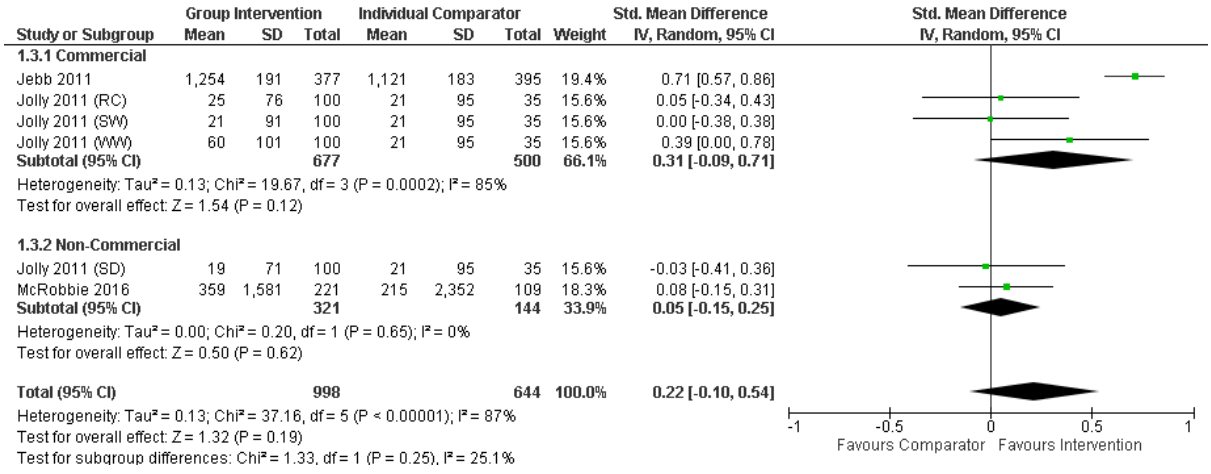
Figure 6: Achievement of 5% Weight Loss at Programme-end (All Studies)



3.5.2 Physical Activity

In total, 3 studies reported outcomes on physical activity levels at 1-year (Figure 7). The amount of time spent being physically active was a SMD 0.22 greater (95% CI -0.10, 0.54; p= 0.19) in group compared to individual interventions. The SMD represents a small effect size given that is less than 0.50 (Cohen 1988). The high heterogeneity in these results was not dealt with by group provider sub-group analysis (p= 0.25).

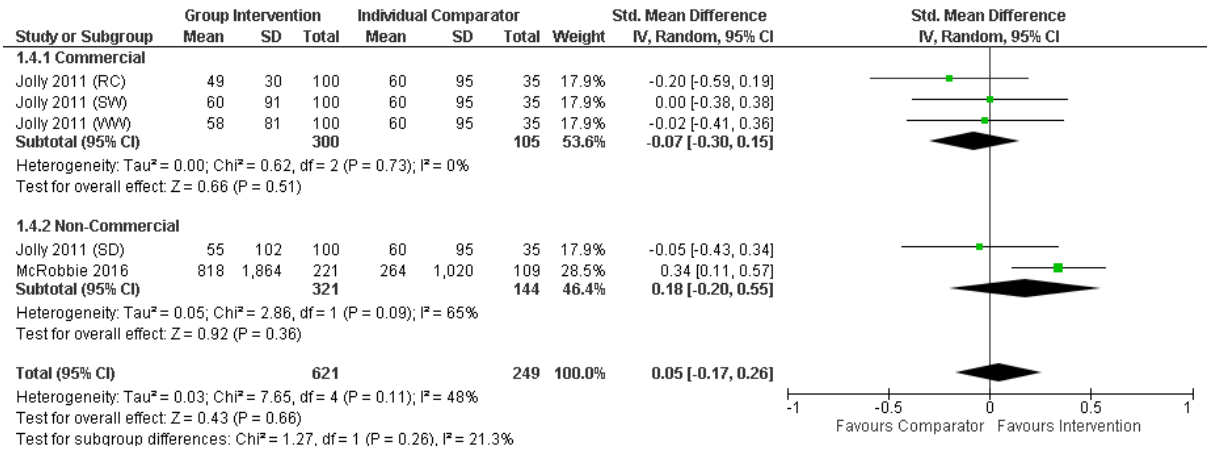
Figure 7: Change in Time Spent being Physically Active at 1-year (All Studies)



The forest plot shown in Figure 8 illustrates results from 2 studies at the programme-end time point. Neither group nor individual intervention was favourable for physical activity levels (SMD 0.035 95% CI [-0.17, 0.26], p= 0.66), based on moderately heterogeneous results (I^2 48%, p= 0.11). Sub-group analysis by provider did not explain heterogeneity (p= 0.26). A sensitivity analysis, based on risk of bias, was not performed because all studies were deemed to be at a low risk.

Based on these results, commercial, non-commercial nor individual intervention were clinically favourable to changes in physical activity levels in the longer- (1-year) or shorter-(programme-end) term.

Figure 8: Change in Time Spent being Physically Active at Programme-end (All Studies)



3.5.3 Quality of Life

Quality of life scores were reported by a single non-commercial intervention (Appel et al. 2011) at the 1-year time-point only. EQ-5D questionnaire scores were reported on the single index and as a visual analogue scale (VAS). The change in single index score was very marginal, with a 0.0004 increase (95% CI -0.03 to 0.03; p= >0.05) in the non-commercial group over the individual intervention. There was a 2.35 point increase (95% CI -2.07 to 6.78) on the visual analogue scale in the non-commercial group compared with individual intervention. This was not statistically significant (p= >0.05) or clinically meaningful, given that the visual analogue scale is measured on a scale of 0 to 100 (Herdman et al. 2011). Therefore, there was no benefit in attending either a non-commercial group or an individual intervention. No determination can be made about commercially provided studies due to the lack of reported data.

3.5.4 Dietary Intake

Dietary intake was assessed by a single commercial group intervention (Jebb et al. 2011); at the 1-year time point only. There were significant favourable changes in nutritional intake in the commercial group over the individual intervention for energy intake (-181kcal, SE 81kcal, p=0.028), total fat intake (-9.6g, SE 4.5g, p= 0.033) and fibre intake (2.6g, SE 0.9g, p=0.016).

Given that most participants were female and middle-aged, these findings were thus compared against relevant dietary reference values (DRVs) (Department of Health 1991) and confirms that these findings are clinically meaningful (see

Table 7).

Measure	n=	Difference in mean	SE	p value	DRV	Difference in DRV[&]
Energy (kcal)	236	-181	81	0.028	2103	-8.6%
Fat (g)	236	-9.6	4.5	0.033	82	-11.7%
Carbohydrates (g)	236	-14.0	9.8	0.153	263	-5.3%
of which Sugars (g)	236	-5.2	6.8	0.445	26	-20%
Fibre Density (g)	236	2.6	0.1	0.016	30	+8.7%

Table 7: Dietary Intake Compared to DRVs

Measure	n=	Difference in mean	SE	p value	DRV	Difference in DRV[&]
Energy (kcal)	236	-181	81	0.028	2103	-8.6%
Fat (g)	236	-9.6	4.5	0.033	82	-11.7%
Carbohydrates (g)	236	-14.0	9.8	0.153	263	-5.3%
of which Sugars (g)	236	-5.2	6.8	0.445	26	-20%
Fibre Density (g)	236	2.6	0.1	0.016	30	+8.7%

[&]Dietary Reference Values for 35 – 54-year-old females

Although carbohydrate and sugar intake favourably decreased, this was not statistically significant (-14.0g, SE 9.8g, p= 0.153; -5.2g, SE 6.8g, p=0.445 respectively). However, when comparing the difference in sugar intake between interventions to the DRV, it is arguable that the sugar intake reductions are in fact clinically significant; having reduced sugar intake by a proportional 20% of the DRV (

Measure	n=	Difference in mean	SE	p value	DRV	Difference in DRV[&]
Energy (kcal)	236	-181	81	0.028	2103	-8.6%

Fat (g)	236	-9.6	4.5	0.033	82	-11.7%
Carbohydrates (g)	236	-14.0	9.8	0.153	263	-5.3%
of which Sugars (g)	236	-5.2	6.8	0.445	26	-20%
Fibre Density (g)	236	2.6	0.1	0.016	30	+8.7%

Table 7)

These results show that commercially provided group interventions are favoured over individual interventions for beneficial reductions in intake of energy, fat and sugar and a favourable increase for the intake of fibre.

These findings are limited, however, being based on only one study and, further, conclusions cannot be drawn for non-commercially provided group interventions due to the absence of reported data.

3.5.5 Mental Health

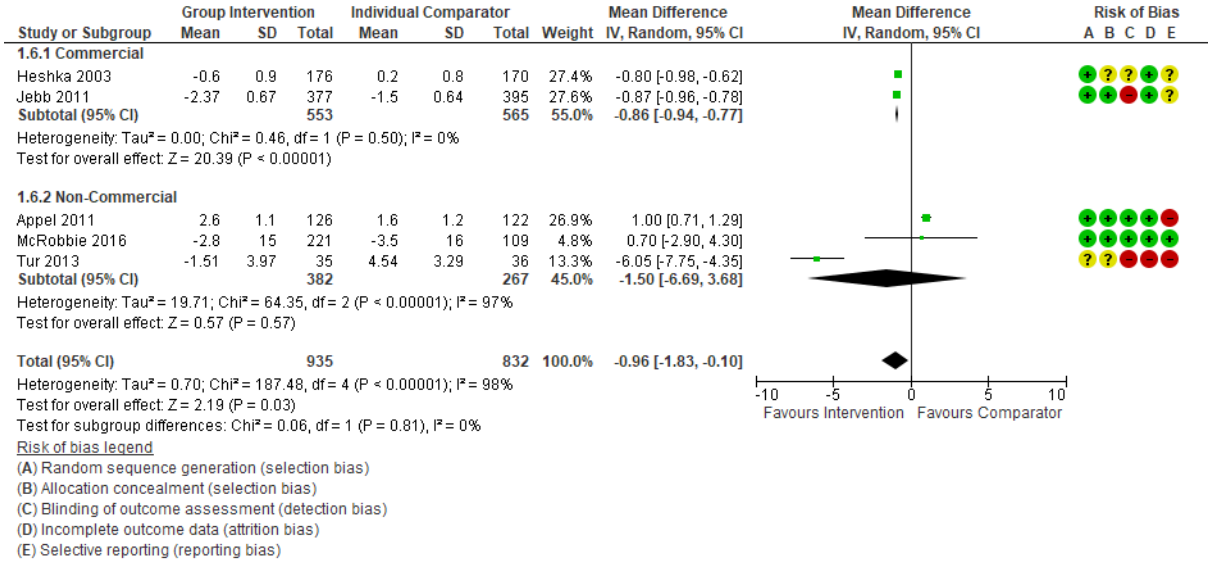
Changes in self-reported measures of mental health were assessed by one study (Appel et al. 2011) at the 1-year time point only. The mean score of depression severity increased by 0.54 points (95% CI -0.27 to 1.36) in the non-commercial group, compared to the individual intervention, which was not significant (p= >0.05). Neither was this considered to be a clinically important score change, given that the PHQ-8 questionnaire is scored on a scale of 0 to 24 points (Kroenke et al. 2009).

Thus, based on the limited available data, neither non-commercial group nor individual intervention demonstrates benefit in self-reported measures of mental health. Conclusions could not be drawn regarding commercially provided group interventions due to the absence of data.

3.5.6 Systolic Blood Pressure

Five studies examined systolic blood pressure at 1-year follow-up. Analysis showed that systolic blood pressure reduced by -0.96mmHg (95% CI -1.83 to -0.10mmHg) and was statistically favourable to group over individual intervention (p= 0.03). There is minimal clinical importance to this change, being less than 1mmHg. What is more, these results were heterogeneous (I^2 98%, p= <0.00001) and sub-group analysis did not explain heterogeneity (p= 0.81) – see Figure 9.

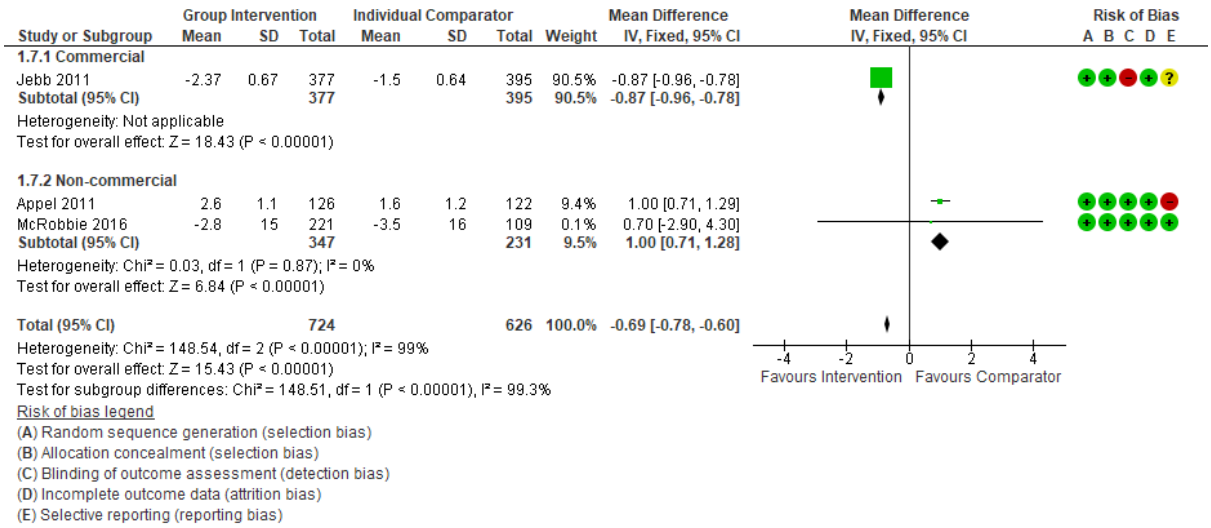
Figure 9: Change in Systolic Blood Pressure at 1-year (All Studies)



Sensitivity analysis found a significant difference, however, between sub-groups (p= <0.00001) when 'high' or 'unclear' risk of bias studies were removed (Figure 10). This

sensitivity analysis found that commercial groups were favoured over individual interventions, with a change in systolic blood pressure of a marginal -0.87mmHg (95% CI -0.96mmHg to -0.78mmHg; p= <0.00001). In contrast, individual interventions were favoured over non-commercial group interventions. Non-commercial group interventions increased blood pressure by a marginal 1.00mmHg (95% CI 0.71mmHg to 1.28mmHg; p= <0.00001). Based on studies deemed to be at a low risk of bias, commercial and non-commercial analyses showed minimal clinically important change in systolic blood pressure.

Figure 10: Change in Systolic Blood Pressure at 1-year (Low Risk of Bias Studies)



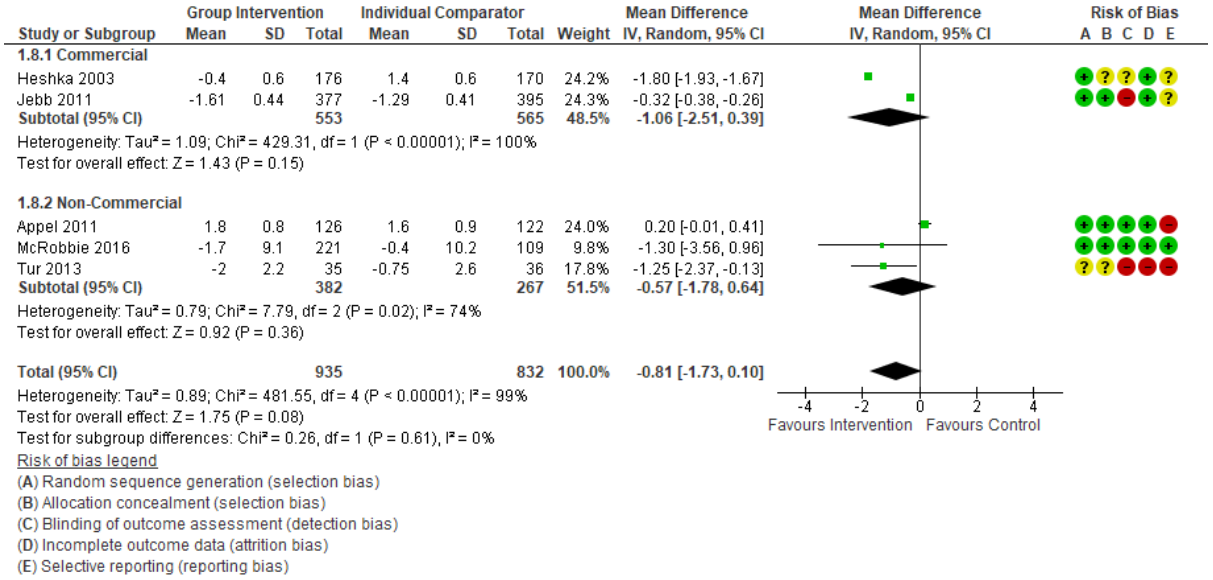
At the programme-end time-point, only one study reported on systolic blood pressure (McRobbie et al., 2016) and therefore meta-analysis was not performed. Systolic blood pressure increased significantly by 5.6mmHg (95% CI 1.0mmHg to 10.3mmHg; p= 0.02) in the group compared to individual intervention. Given that a systolic blood pressure above 140mmHg is diagnostic for stage 1 hypertension (NICE 2011); 5.6mmHg equates to a 4% proportional increase which is of clinical importance.

Therefore, individual intervention is clinically favoured over a non-commercial group intervention, but is based on limited evidence from one study. Conclusions cannot be drawn about commercially provided group interventions at the programme-end time-point.

3.5.7 Diastolic Blood Pressure

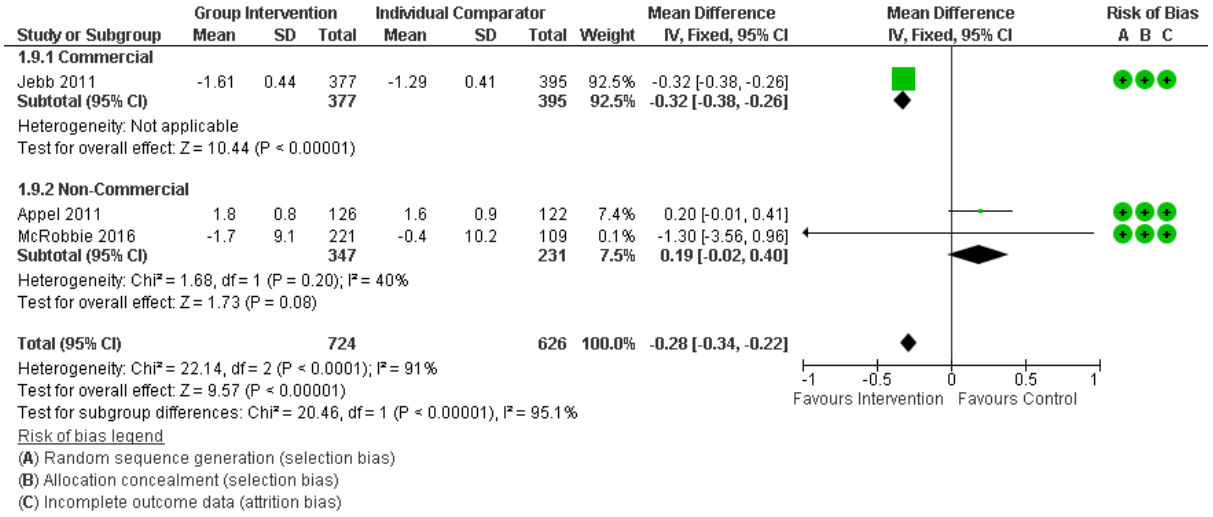
Five studies reported on diastolic blood pressure at 1-year (Figure 11). Blood pressure reduced by a clinically negligible -0.81mmHg (95% CI -1.73mmHg to 0.10mmHg; p= <0.00001) in group compared to individual intervention. Heterogeneity existed (I^2 99%, p= >0.00001) and was not explained by sub-group analysis (p= 0.61).

Figure 11: Change in Diastolic Blood Pressure at 1-year (All Studies)



Performing a sensitivity analysis by including only the low risk of bias studies (Figure 12) did explain differences between sub-groups (p= <0.00001) and thus dealt with heterogeneity. Commercial groups were favoured over individual interventions (p= <0.00001) but this was not deemed to be clinically meaningful based on a reduction in blood pressure of 0.32mmHg (95% CI -0.38mmHg to -0.26mmHg). Meanwhile, neither non-commercial groups nor individual interventions were favoured based on an increase in blood pressure of 0.19mmHg (95% CI -0.02mmHg to 0.40mmHg; p= 0.08).

Figure 12: Change in Diastolic Blood Pressure at 1-year (Low Risk of Bias Studies)



Diastolic blood pressure at the programme-end was measured only by McRobbie et al. (2016), a non-commercial group intervention. Their results did not favour either the group or individual intervention (p= 0.81), with a reduction in blood pressure of 0.3mmHg (95% CI -3.0mmHg to 2.4mmHg). Meanwhile, conclusions could not be drawn for commercial group interventions due to lack of data.

3.5.8 Lipid Profile

Overall, five meta-analyses were performed to analyse the five different reported measures of changes to lipid profile levels at 1-year: total cholesterol, total-to-HDL cholesterol ratio, HDL cholesterol, LDL cholesterol and triglycerides. No studies reported on lipid markers at the programme-end time point.

3.5.8.1 Total Cholesterol

Group intervention was found to reduce levels of total cholesterol by -0.03mmol/L (95% CI -0.03mmol/L to -0.02mmol/L; p= <0.00001) based on 4 studies and findings were

heterogeneous (Figure 13). Sub-group analysis by group provider showed differences between sub-groups (p=<0.00001) which dealt with heterogeneity.

Group intervention was statistically favoured when provided commercially, reducing cholesterol by -0.03mmol/L (95% CI -0.03mmol/L to -0.02mmol/L; p= <0.00001); albeit the clinical significance of this is negligible. Meanwhile, total cholesterol increased in the non-commercial group intervention compared to individual intervention by 2.06mmol/L (95% CI 1.82mmol/L to 2.30mmol/L; p= <0.00001) and therefore individual intervention is favoured. This is judged to be clinically significant given that this equates proportionally to a change of over half (52%) of the 4.0mmol/L optimal total cholesterol target (British Cardiac Society et al. 2005). A sensitivity analysis, including only 'low' risk of bias studies, did not alter the effect sizes or homogeneity for either sub-group (Figure 14).

Figure 13: Change in Total Cholesterol at 1-year (All Studies)

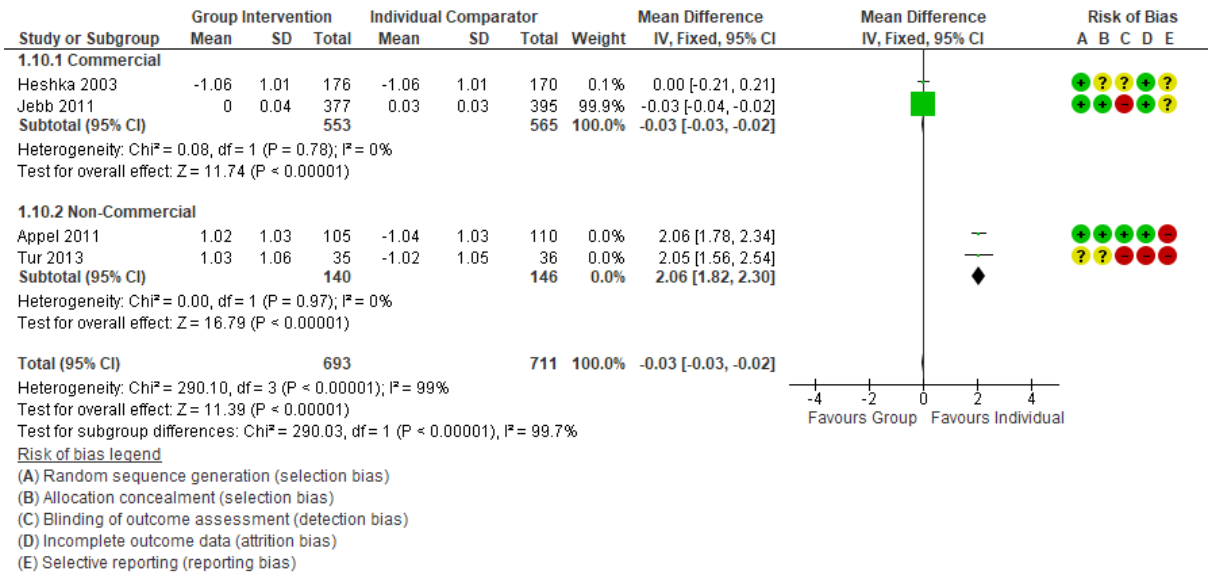
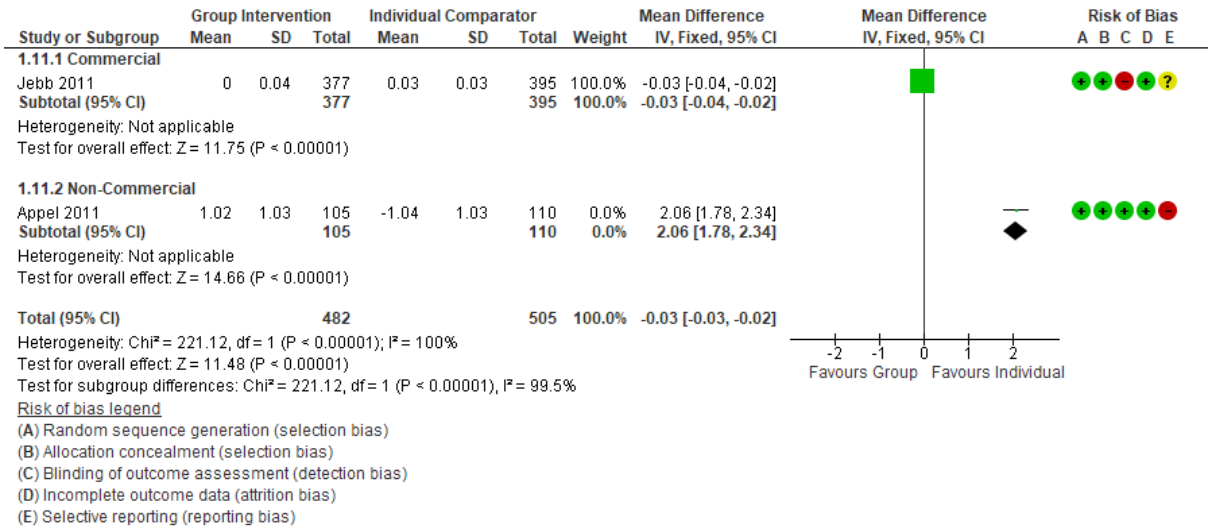


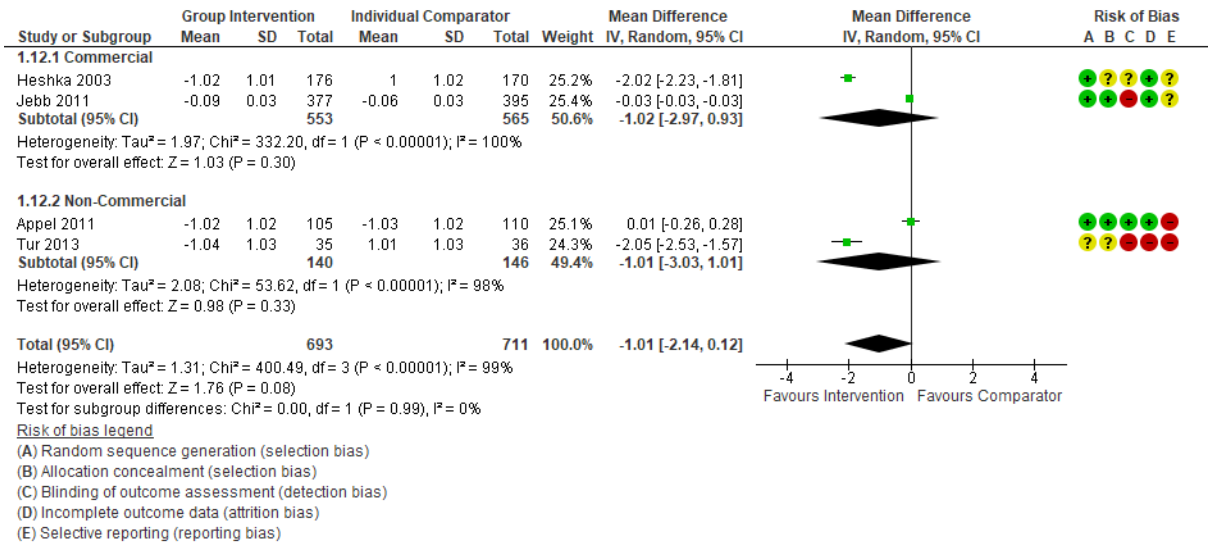
Figure 14: Change in Total Cholesterol at 1-year (Low Risk of Bias Studies)



3.5.8.2 Triglycerides

At 1-year, neither group nor individual intervention was favoured according to triglyceride changes (-1.01mmol/L, 95% CI [-2.14, 0.12]; p= 0.08) and results were heterogeneous (I^2 99%; p= <0.00001). Sub-group analysis by group provider did not explain any differences between groups (p= 0.99) and thus did not address heterogeneity (Figure 15).

Figure 15: Change in Triglycerides at 1-year (All Studies)

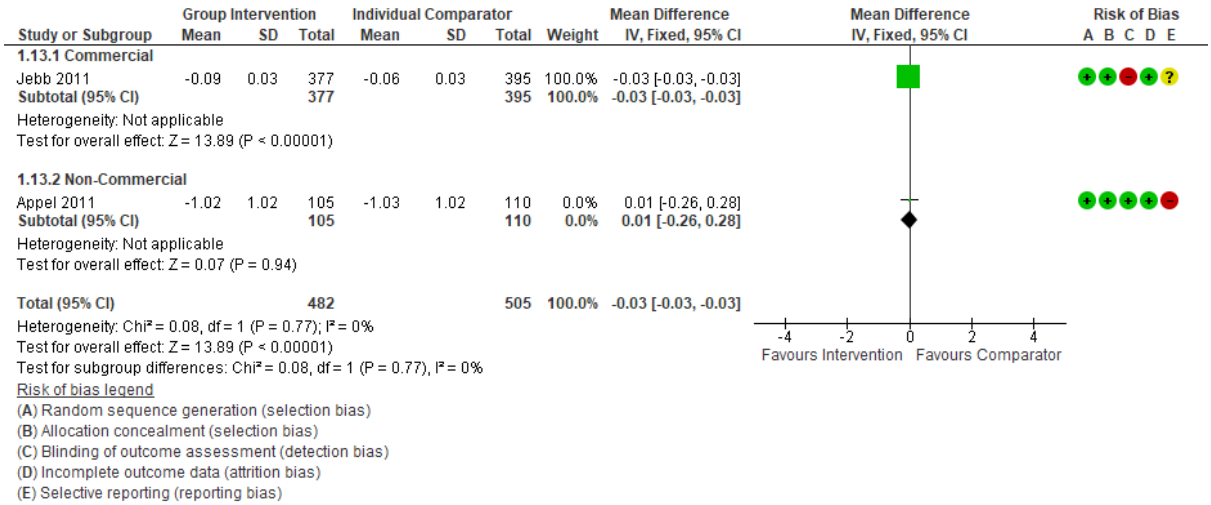


Sensitivity analysis including only low risk of bias studies did address heterogeneity; but in turn also influenced the results of the meta-analysis (Figure 16).

Results from low risk of bias studies were homogenous (I^2 0%, p= 0.77). These results found that triglyceride levels reduced by -0.03mmol/L (95% CI -0.03mmol/L to -0.03mmol/L; p= <0.00001) compared to individual interventions. In the clinical context, however, this is a clinically negligible change. Proportionally this equates to a 2% change when compared with the >1.7mmol/L diagnostic for elevated triglycerides (British Cardiac Society et al. 2005).

Therefore, commercial groups, non-commercial groups nor individual interventions were clinically favoured in relation to changes in triglyceride levels; based on low risk of bias studies.

Figure 16: Change in Triglycerides at 1-year (Low Risk of Bias Studies)

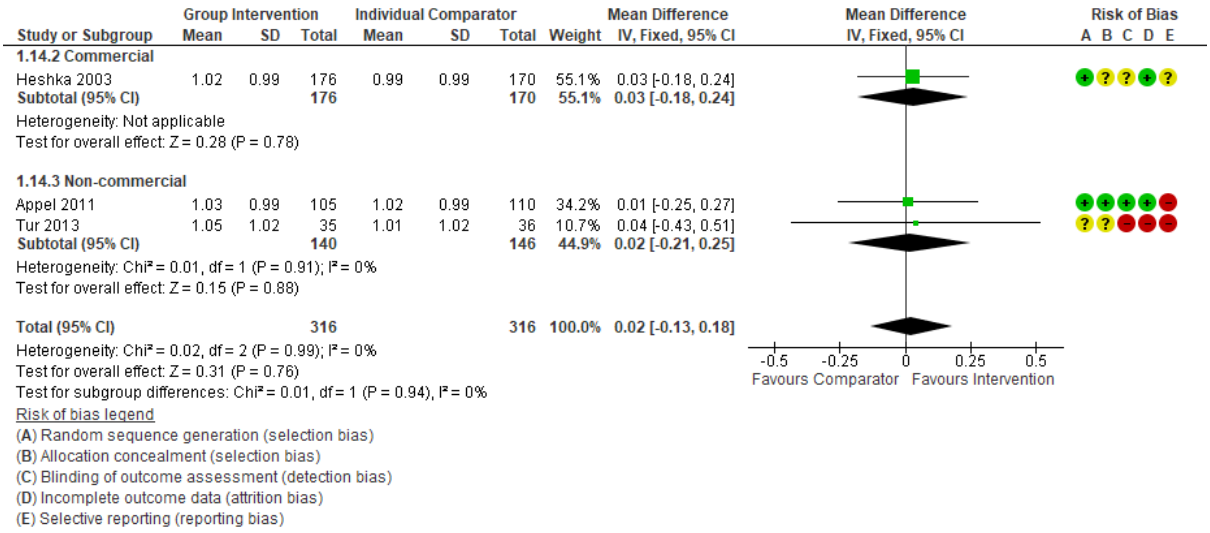


3.5.8.3 HDL Cholesterol

Neither group nor individual intervention favoured changes in HDL cholesterol (0.02mmol/L, 95% CI [-0.13, 0.18], p= 0.76) and results were homogenous (I2 0%, p= 0.99) (Figure 17). Sub-group analysis by provider showed no differences between commercial and non-commercial group interventions (p= 0.94). Thus, commercial, non-commercial nor individual

interventions were favourable in relation to beneficial changes in HDL cholesterol levels. This finding was not altered by the risk of bias of the included studies as results were entirely homogenous.

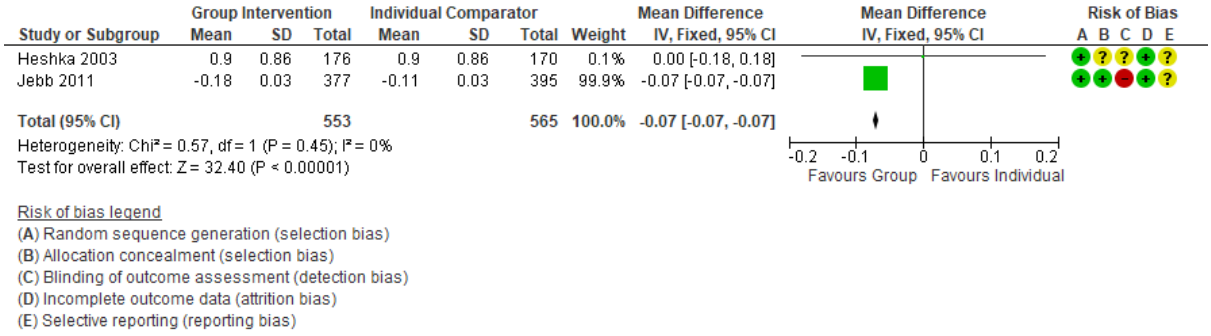
Figure 17: Change in HDL Cholesterol at 1-year (All Studies)



3.5.8.4 Total-to-HDL Cholesterol Ratio

Total to HDL cholesterol ratio was reported by studies exclusively using commercial group providers. Results were homogenous (I^2 0%, p= 0.45), regardless of risk of bias, and favoured group over individual intervention (-0.07, 95% CI [-0.07, -0.07], p= <0.00001) (Figure 18). The clinical benefit of this effect is negligible when put into context that a ratio of <4.0 is beneficial for cardiovascular health (British Cardiac Society et al. 2005). Based on this, neither a commercial group nor an individual intervention is favoured in improving total-to-HDL cholesterol ratio. No conclusions can be drawn about non-commercial groups due to the lack of data from such providers.

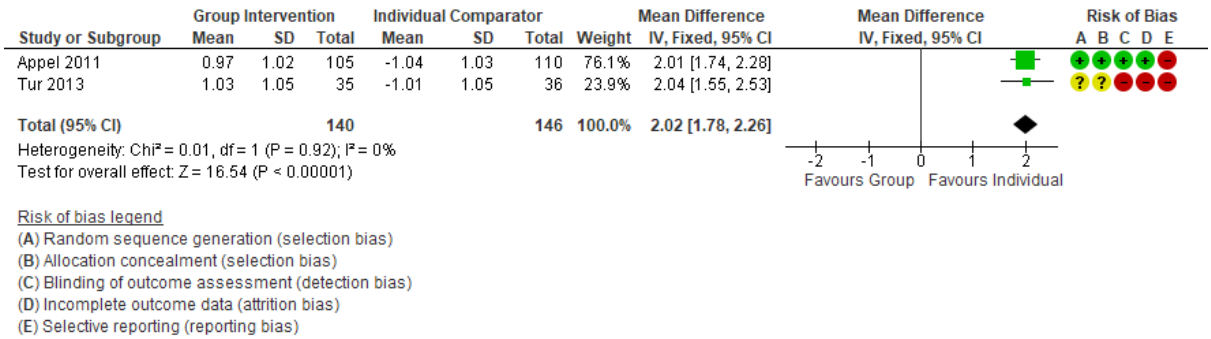
Figure 18: Change in Total-to-HDL Cholesterol Ratio at 1-year (All Studies)



3.5.8.5 LDL Cholesterol

LDL cholesterol was measured in 2 studies which exclusively used non-commercial group providers (Figure 19). There were increased levels of LDL cholesterol in non-commercial groups (2.02mmol/L, 95% CI [1.78mmol/L, 2.26mmol/L], p= <0.00001) compared to individual intervention and results were homogenous (I^2 0%, p= 0.92), regardless of the included studies' risk of bias (Figure 19). These results indicate that individual intervention is clinically favoured over non-commercial group interventions in relation to changes in LDL cholesterol; when considered that the reference range for elevated LDL is above 2.0mmol/L, (British Cardiac Society et al. 2005). Due to the absence of data from commercial group interventions, conclusions on the effectiveness of group interventions from commercial providers cannot be drawn.

Figure 19: Change in LDL Cholesterol at 1-year: All Studies (All Studies)



3.5.9 Glycaemic Markers

Three meta-analyses were performed to analyse the three different reported measures of diabetes disease risk at 1-year: fasting glucose, fasting insulin and HbA1c. No studies reported on glycaemic markers at the programme-end time point.

3.5.9.1 Fasting Glucose

Four studies measured fasting glucose at the 1-year time point. Group intervention was favoured over individual intervention (-0.05mmol/L, 95% CI [-0.05, -0.05], p= <0.00001) but this analysis was heterogeneous (I^2 100%, p= <0.00001). Although there was a significant difference between the sub-groups (p= <0.00001), high levels of heterogeneity remained in both the commercial (I^2 100%, p= <0.00001) and non-commercial (I^2 95%, p= <0.00001) sub-groups (Figure 20). Low risk of bias sensitivity analysis did not alter these findings (Figure 21).

While these findings were statistically significant, clinically these findings have little implication when it is put into context that impaired fasting glucose is determined at a level of more than 7.0mmol/L (NICE 2015). Therefore commercial group, non-commercial group nor individual interventions are favoured in relation to changes in fasting glucose levels.

Figure 20: Change in Fasting Glucose at 1-year (All Studies)

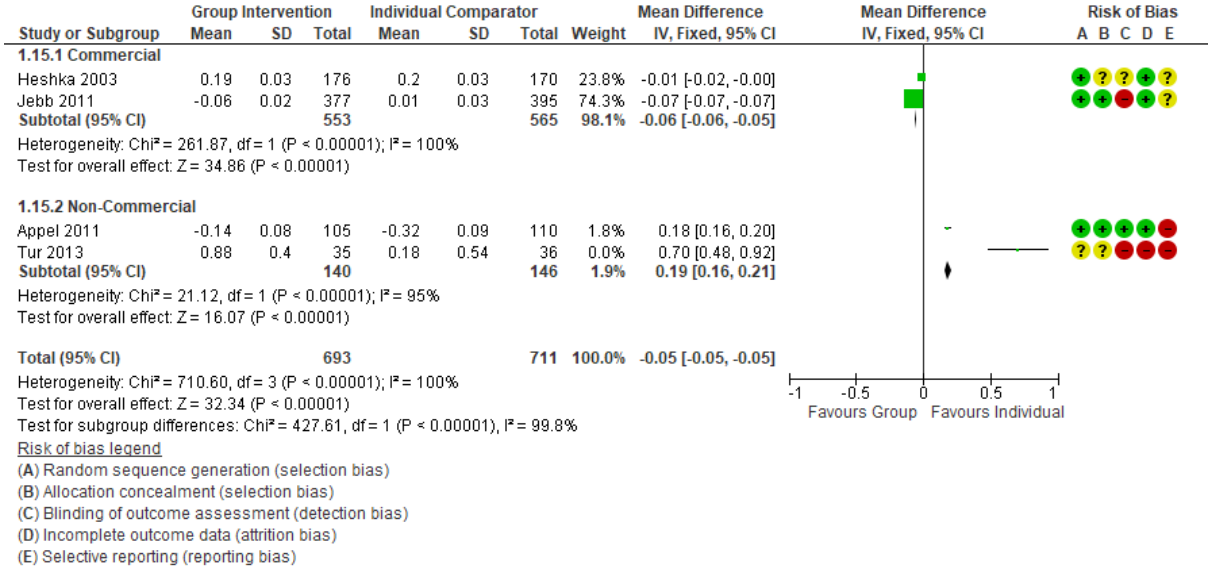
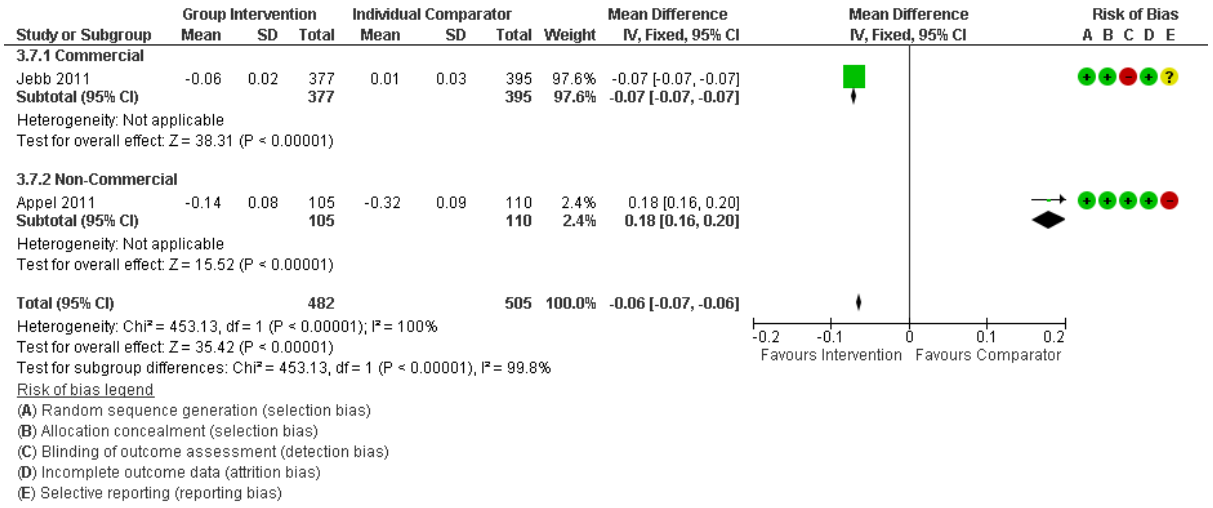


Figure 21: Change in Fasting Glucose at 1-year (Low Risk of Bias Studies)

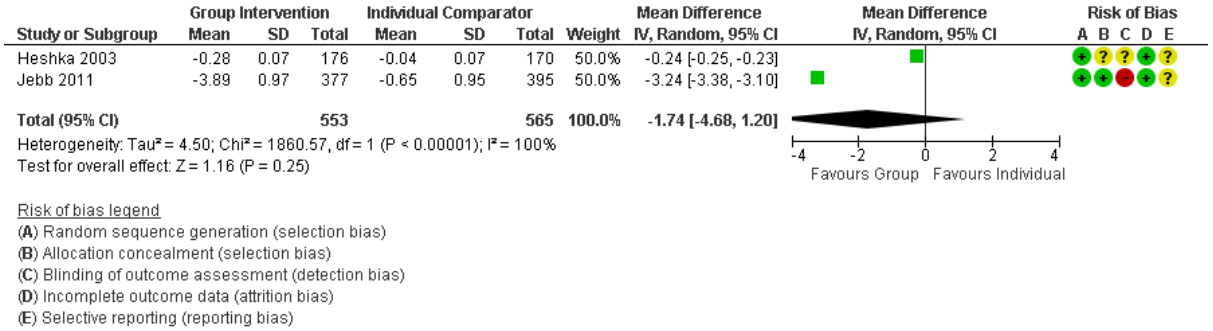


3.5.9.2 Fasting Insulin

Two studies, both using commercial group interventions, reported data on fasting insulin level changes. Neither commercial groups nor individual interventions favoured changes in fasting insulin (-1.74pmol/L, 95% CI [-4.68, 1.20], p= 0.25) and results were heterogeneous (I^2 100%;

p= <0.00001) (Figure 22). The group interventions in both included studies were provided commercially and therefore a sub-group analysis was not performed.

Figure 22: Change in Fasting Insulin at –year (All Studies)

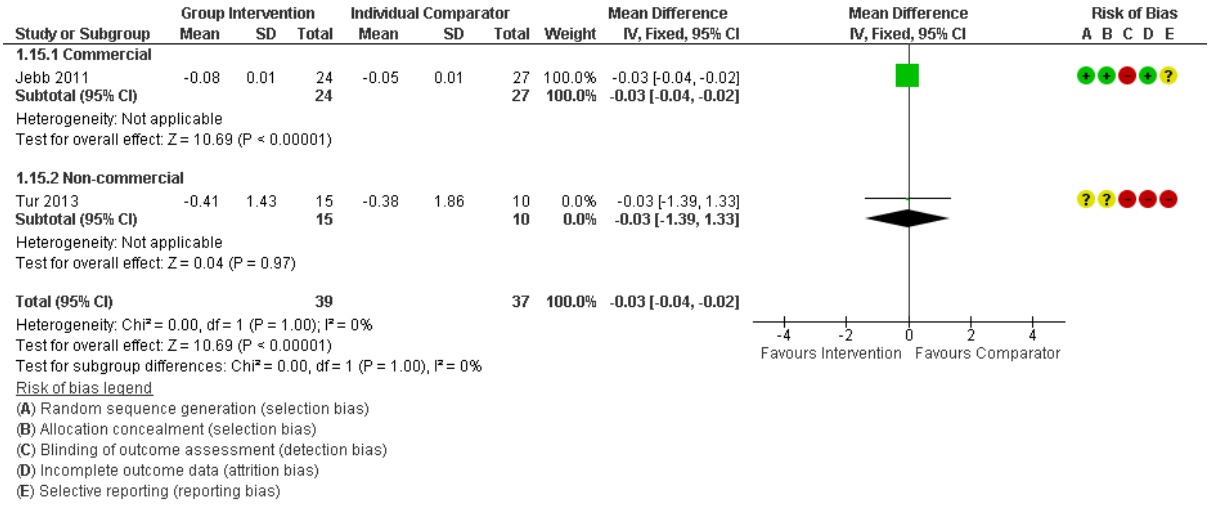


Jebb et al. (2011) was the only 'low' risk of bias study, which on visual inspection has a much larger effect size on fasting insulin levels (-3.24pmol/L, 95% CI [-3.38, -3.10]). However, considering a fasting insulin result greater than 175pmol/L is considered elevated (Melmed and Williams 2011), the clinical importance of this change in fasting insulin is negligible. Therefore, based on one low risk of bias study, the findings are also in agreement that neither a commercial group nor an individual intervention is favoured in relation to changes in fasting insulin levels.

3.5.9.3 HbA1c

Two studies reported changes in HbA1c levels (Figure 23). Any group intervention was favoured over individual intervention, with statistically significant reductions in HbA1c (-0.03mmol/mol, 95% CI [-0.04, -0.02], p= <0.00001) based on homogenous results (I^2 0%, p= 1.00). This was not clinically significant, however, given the marginal change in HbA1c level, when compared to the diagnostic value of 42mmol/mol given for impaired glucose tolerance (NICE 2015). Neither the sub-group nor the sensitivity analyses altered these findings.

Figure 23: Change in HbA1c at 1-year (All Studies)



3.5.10 Attrition

At 1-year follow-up, all 8 studies reported attrition rates. The RR of a participant attending their follow-up visit was favoured by neither group nor individual intervention (RR 1.00, 95% CI [0.95, 1.06], p= 0.93) and results were heterogeneous (I^2 50%, p= 0.02) (Figure 24).

Sub-group analysis by group provider dealt with heterogeneity, finding a significant difference between the results of commercial and non-commercial group interventions (p= 0.002) (Figure 24). Attendance at 1-year follow-up was favoured by neither commercial group nor individual intervention (RR 1.07, 95% CI [0.99, 1.16], p= 0.07). On the other hand, attendance was 7% less likely in a non-commercial group relative to an individual intervention (RR 0.91, 95% CI [0.85 0.98], p= 0.01).

Performing sensitivity analysis, by including only low risk of bias studies, did alter the significance of these findings (Figure 25). Based on low risk of bias studies only, attendance was 14% more likely in commercial groups relative to individual interventions (RR 1.14, 95% CI [1.03, 1.25], p= 0.01). Meanwhile, neither non-commercial or individual interventions were favoured for attendance (RR 0.97, 95% CI [0.90, 1.04], p= 0.34).

Figure 24: Attendance at 1-year (All Studies)

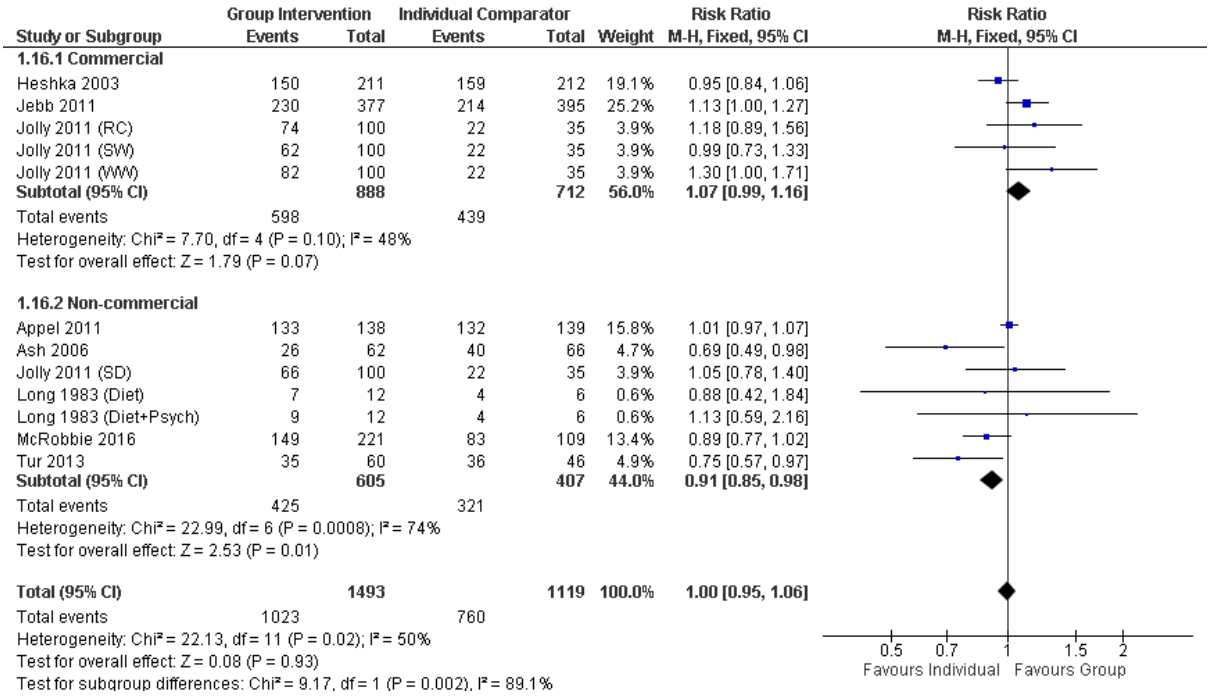
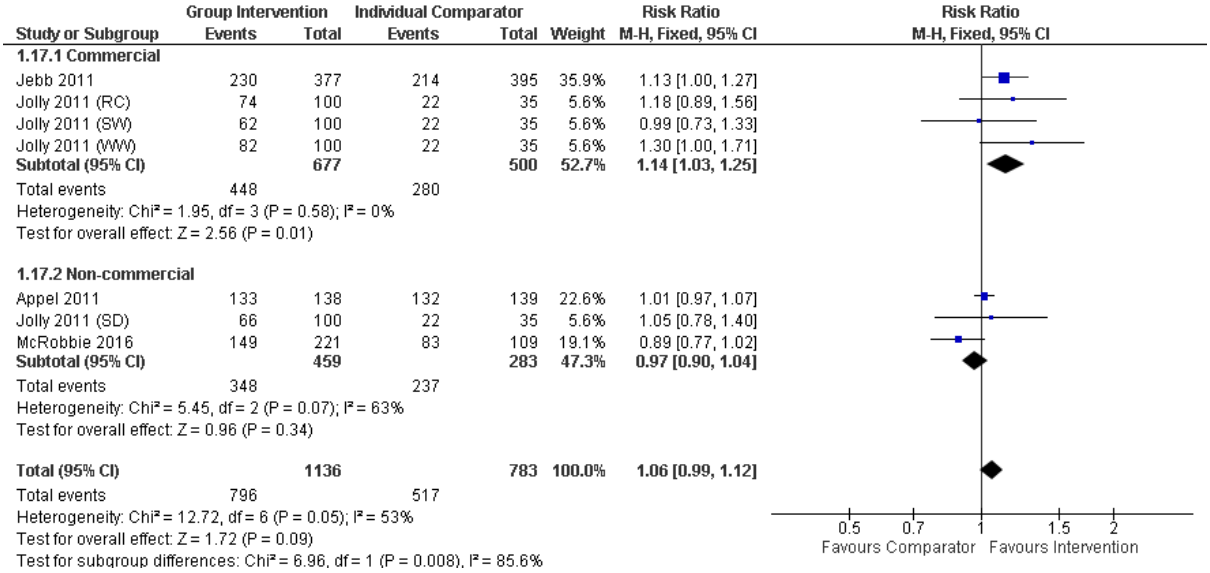


Figure 25: Attendance at 1-year (Low Risk of Bias Studies)



At the programme-end time point, 5 studies reported attrition rates. The RR of attending at programme-end was favoured by neither group or individual intervention (RR 1.04, 95% CI [0.98, 1.09], p= 0.18) (Figure 26).

Sub-group analysis by provider explained a significant difference between the results of commercial and non-commercial groups (p= 0.02) and dealt with heterogeneity (Figure 26). The RR of attendance favoured commercial group over individual intervention (RR 1.10, 95% CI [1.04, 1.16], p= 0.0008), but favoured neither non-commercial nor individual intervention (RR 0.94, 95% CI [0.83, 1.05], p= 0.27) at programme-end (Figure 26). Clinically, this is meaningful given that commercial group participants were 10% more likely to attend at programme-end, relative to individual intervention. Neither non-commercial group nor individual intervention participants were at a greater likelihood of attending at the programme-end. Including only ‘low’ risk of bias studies in the sensitivity analysis did not influence these findings at programme-end (Figure 27).

Figure 26: Attendance at Programme-end (All Studies)

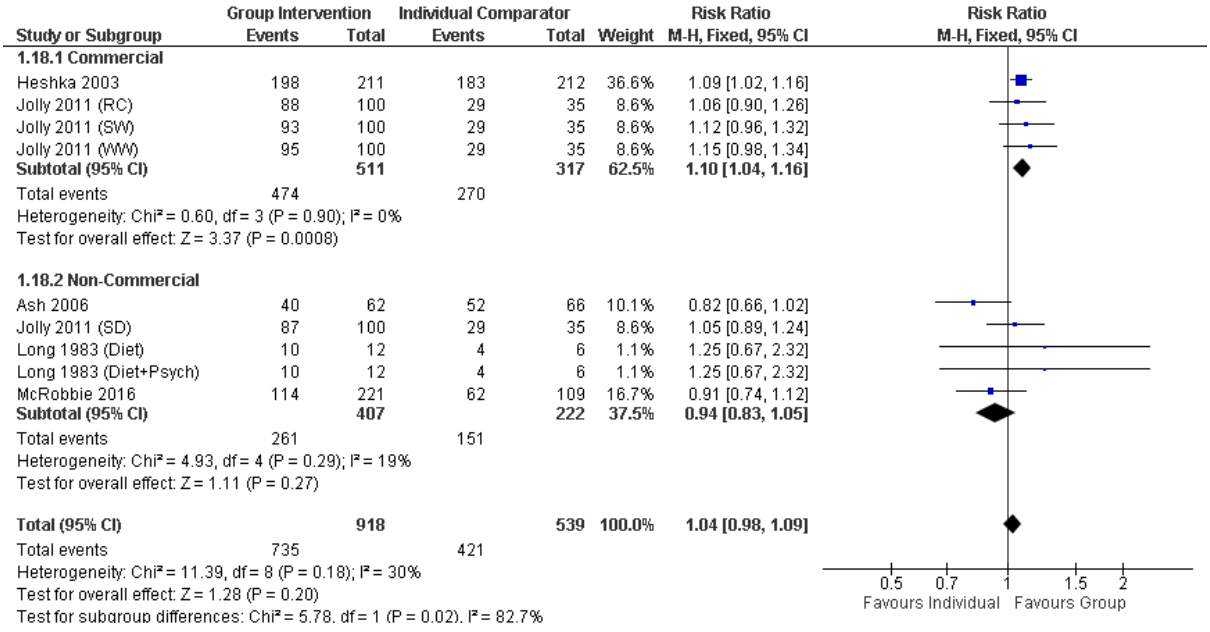
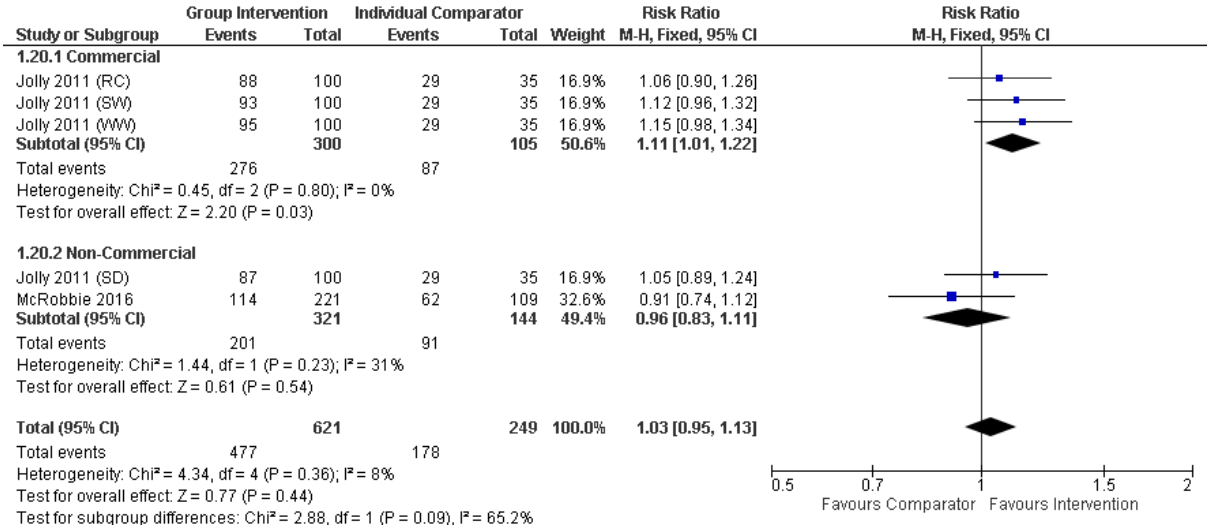


Figure 27: Attendance at Programme-end (Low Risk of Bias Studies)



3.5.11 Related Adverse Events

The incidence of adverse events that may be related to the study interventions were reported as 1/662 (0.2%) in the group intervention and 0/687 in the individual intervention (0.0%). The SAE for this 1 participant was detailed by the authors to be a hospitalisation following musculoskeletal injuries. These injuries were obtained following an assault whilst exercising outdoors. Due to the rarity in events, meta-analysis was not possible.

3.5.12 Cost-effectiveness

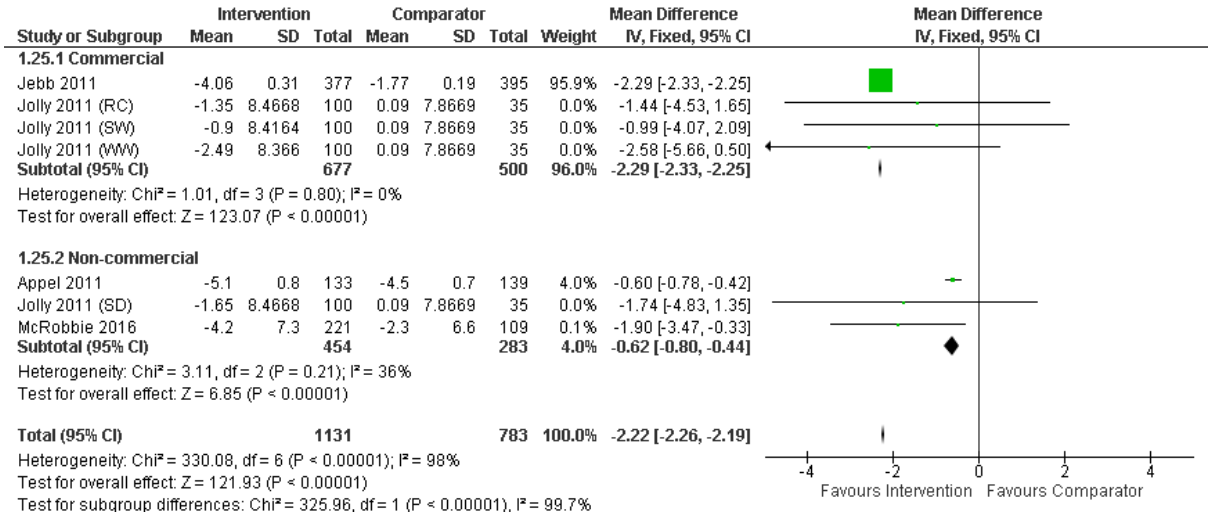
Cost-effectiveness was reported by both Jebb et al. (2011) and McRobbie et al. (2016) as cost per QALY. In the UK, cost per QALY was far lower than the £20,000 - £30,000 NICE threshold and was between £8,128 (Jebb et al. 2011) to £8,439 (McRobbie et al. 2016). The costs per QALY were greater in other healthcare systems in Australia (£12,134) and in Germany (£14,416) as reported by Jebb et al. (2011). These costs are based upon imputations for the financial year of 2016 (see Appendix 5).

3.6 Additional Sensitivity Analyses

It was identified in section 3.4 that not all included studies in this review reported dichotomous data on achievement of 5% weight loss; however, 7 of the included studies had extractable continuous data available on absolute weight loss, measured in kilograms, at the 1-year time point.

Results from these sensitivity analyses showed that, based on low risk of bias studies, groups were favoured over individual interventions; but results were highly heterogeneous (-2.22kg, 95% CI [-2.26, -2.19], p= <0.00001; I^2 98%) (Figure 28). By sub-group, commercial and non-commercial group interventions were also favoured over individual interventions. Results were homogenous for commercial groups (-2.29kg, 95% CI [-2.33, -2.25], p= <0.00001; I^2 0%) but heterogeneous for non-commercial groups (-0.62kg, 95% CI [-0.80, -0.44], p= 0.21; I^2 36%).

Figure 28: Sensitivity Analysis of Weight Change (kg) at 1-year (Low Risk of Bias Studies)



Thus, sensitivity analysis confirmed that measurement of weight loss as a continuous outcome measure, in kilograms, did not change the direction, significance or magnitude of effect at the 1-year time point, based on usable data for 7 of the 8 studies.

3.7 Publication Bias

All studies reported a statistically significant p-value (<0.05) for the pre-defined primary outcome measure of weight change. Considering the absence of non-significant p-values that only one study had a smaller sample sizes (<100 participants), it is plausible that there may exist a potential publication bias within this review. It could be that publication is favoured by larger, and thus more substantially funded, studies. Moreover, it may be conceivable that studies with a non-significant statistical result were not published.

Table 8: Significance of Primary Outcome by Sample Size

Sample size	Statistical Significance (p= <0.05)		Total
	Significant	Not significant	
≤20	-	-	0
21-100	1	-	1
>100	7	0	7

3.8 Quality of Evidence (GRADE)

Six outcomes were selected to be critical or important at the 1-year time point. These outcomes included the likelihood of a participant achieving a 5% weight loss, changes in systolic blood pressure, total to HDL cholesterol ratio, HbA1c and quality of life; as well cost-effectiveness. Results of these can be found in

	Certainty assessment								
Provider	№ of participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event rates (%) With Individual	With Group
Achievement of greater than 5% weight loss									
Any	1920 (5 RCTs)	not serious	serious [a]	not serious	not serious	None	⊕⊕⊕◯ MODERATE	196/764 (25.7%)	430/1156 (41.6%)
Commercial	1191 (3 RCTs)	not serious	not serious	not serious	not serious	None	⊕⊕⊕⊕ HIGH	119/489 (24.3%)	293/702 (46.0%)
Non-commercial	729 (3 RCTs)	not serious	not serious	not serious	serious [f]	None	⊕⊕⊕◯ MODERATE	77/275 (28.0%)	137/454 (33.9%)
Changes in Systolic Blood Pressure									
Any	1350 (3 RCTs)	not serious	very serious [a, B]	serious [c]	serious [f]	None	⊕◯◯◯ VERY LOW	626	724
Commercial	772 (1 RCT)	not serious	not serious	serious [c]	not serious	None	⊕⊕⊕◯ MODERATE	395	377
Non-commercial	578 (2 RCTs)	not serious	not serious	serious [c]	not serious	None	⊕⊕⊕◯ MODERATE	231	347
Changes in Total:HDL Cholesterol									
Any	1118 (2 RCTs)	not serious	not serious	serious [c]	not serious	None	⊕⊕⊕◯ MODERATE	565	553
Changes in HbA1c									
Any	76 (2 RCTs)	not serious	not serious	not serious	serious [d]	None	⊕⊕⊕◯ MODERATE	37	39
Changes in Quality of Life									
Non-commercial	277 (1 RCT)	not serious	not serious	serious [c]	very serious [e, f]	None	⊕◯◯◯ VERY LOW	MD 0.0004 EQ-5D score (	
Cost-effectiveness									
Any	1102 (2 RCTs)	not serious	not serious	not serious	not serious	None	⊕⊕⊕⊕ HIGH	ICER for group interventi 2016 finan	

3.8.1 Achieving a 5% weight loss

The certainty in the evidence for commercial and non-commercial groups combined was 'moderate'. A downgrade was applied for inconsistency due to heterogeneity (I^2 35%). Meanwhile, the certainty of evidence for non-commercial groups alone was also 'moderate' and was downgraded for imprecision due to a confidence interval which spanned a favourable and non-favourable effect. The certainty of the evidence for commercial groups was 'high' as no downgrade was applied. In-directly comparing commercial to non-commercial group interventions showed clinically very important findings based on 'moderate' to 'high' quality evidence. Out of 1,000 participants, 217 more participants would lose 5% weight in a commercial group programme relative to an individual intervention. Meanwhile, 59 more

participants out of 1,000 would lose 5% weight loss in a non-commercial programme, relative to an individual intervention.

3.8.2 Systolic Blood Pressure

Changes in systolic blood pressure were clinically marginal and thus neither group nor individual intervention was favoured regarding changes in systolic blood pressure. Several downgrades were applied to the quality of the evidence and was deemed to be of ‘very low’ quality. Inconsistency was downgraded twice, being ‘very serious’, due to heterogeneity being confirmed by both I^2 (99%) and p-value (<0.00001). The certainty of indirectness was downgraded to ‘serious’ because of this measure being a surrogate marker for risk of cardiovascular disease. There was also ‘serious’ imprecision because of the wide interval of confidence of effect size.

Table 9: GRADE Assessment and Summary of Findings

Provider	Certainty assessment							Summary of findings					
	№ of participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects		Number needed to treat (95% CI)
								With Individual	With Group		Risk with Individual	Risk difference with Group	
Achievement of greater than 5% weight loss													
Any	1920 (5 RCTs)	not serious	serious [a]	not serious	not serious	None	⨁⨁⨁◯ MODERATE	196/764 (25.7%)	430/1156 (41.6%)	RR 1.62 (1.40 to 1.86)	257 per 1,000	159 more per 1,000 (103 more to 221 more)	9 (6.4 to 13.6)
Commercial	1191 (3 RCTs)	not serious	not serious	not serious	not serious	None	⨁⨁⨁⨁ HIGH	119/489 (24.3%)	293/702 (46.0%)	RR 1.89 (1.58 to 2.26)	243 per 1,000	217 more per 1,000 (141 more to 307 more)	6 (4.4 to 8.2)
Non-commercial	729 (3 RCTs)	not serious	not serious	not serious	serious [f]	None	⨁⨁⨁◯ MODERATE	77/275 (28.0%)	137/454 (33.9%)	RR 1.21 (0.95 to 1.54)	280 per 1,000	59 more per 1,000 (14 fewer to 151 more)	46 (N/A)
Changes in Systolic Blood Pressure													
Any	1350 (3 RCTs)	not serious	very serious [a, B]	serious [c]	serious [f]	None	⨁◯◯◯ VERY LOW	626	724	-	-	MD 0.15 mmHg higher (1.53 lower to 1.84 higher)	-
Commercial	772 (1 RCT)	not serious	not serious	serious [c]	not serious	None	⨁⨁⨁◯ MODERATE	395	377	-	-	MD 0.87 lower (0.96 lower to 0.78 lower)	-
Non-commercial	578 (2 RCTs)	not serious	not serious	serious [c]	not serious	None	⨁⨁⨁◯ MODERATE	231	347	-	-	MD 1 higher (0.71 higher to 1.28 higher)	-
Changes in Total:HDL Cholesterol													
Any	1118 (2 RCTs)	not serious	not serious	serious [c]	not serious	None	⨁⨁⨁◯ MODERATE	565	553	-	-	MD 0.07 mmol/L lower (0.07 lower to 0.07 lower)	-
Changes in HbA1c													
Any	76 (2 RCTs)	not serious	not serious	not serious	serious [d]	None	⨁⨁⨁◯ MODERATE	37	39	-	-	MD 0.03%mmol/mol lower (0.04 lower to 0.02 lower)	-
Changes in Quality of Life													
Non-commercial	277 (1 RCT)	not serious	not serious	serious [c]	very serious [e, f]	None	⨁◯◯◯ VERY LOW	MD 0.0004 EQ-5D score (0.03 lower to 0.03 higher). MD 2.35 EQ-5D VAS score (2.07 lower to 6.78 higher)					-
Cost-effectiveness													
Any	1102 (2 RCTs)	not serious	not serious	not serious	not serious	None	⨁⨁⨁⨁ HIGH	ICER for group intervention over individual intervention in the UK according to 2016 financial year calculations is £8,128 - £8,439					-

Explanations: [a] I^2 suggests heterogeneity (greater than0%), [b] p-value suggests heterogeneity (greater than0.05), [c] surrogate marker, [d] very small sample size (n= <100), [e] small sample size. [f] wide confidence interval

When considering evidence from non-commercial group intervention providers and commercial group providers as separate interventional approaches, neither group nor individual intervention from either provider was favoured. This is due to the clinically marginally changes in systolic blood pressure. The quality of evidence for both commercial and non-commercial groups was 'moderate', with one downgrade being applied to both for indirectness. This downgrade was assigned because systolic blood pressure is a surrogate marker of cardiovascular disease risk.

3.8.3 Total to HDL Cholesterol Ratio

The quality of the evidence for this outcome was downgraded to 'moderate' since this is a surrogate marker of cardiovascular risk. Results from commercial and non-commercial groups were included together and thus this evidence is applicable to any group intervention. Changes in total to HDL cholesterol ratio were not clinically superior in group or individual intervention. Based on this 'moderate' quality of evidence, neither group nor individual intervention was favoured in regard to clinical changes in total to HDL cholesterol ratio.

3.8.4 HbA1c

Clinically, changes in HbA1c favoured neither any group nor an individual intervention. The quality of evidence for this outcome is 'moderate', downgraded due to imprecision because of a small sample size of 76 participants.

3.8.5 Quality of Life

The risk of bias for evidence of quality of life is 'low', albeit based upon one study. Several certainty downgrades were applied to this outcome, including one downgrade for the indirectness of questionnaire assessment and two downgrades for imprecision; due a small sample size of n= 277, and confidence intervals that spanned favourable and non-favourable findings for the EQ-5D VAS scores. Clinically, due to marginal changes in quality of life, neither

a non-commercial group nor an individual intervention was favoured based upon the ‘low’ certainty of evidence.

3.8.6 Cost-effectiveness

No downgrades were applied to the quality of evidence for the outcome of cost-effectiveness. There is, therefore, ‘high’ certainty evidence to suggest that both commercial and non-commercial group interventions are equally cost-effective when compared to individual intervention, with a cost per QALY of £8,128 and £8,439 for commercial and non-commercial groups respectively.

4.0 Discussion

4.1 Summary of Key Findings

This systematic review was conducted with the purpose of investigating the efficacy of group relative to individual lifestyle interventions for adult weight management. There is sufficiently robust evidence from important measures (5% weight loss attainment, cost-effectiveness and changes in HbA1c, systolic blood pressure and total:HDL cholesterol) to determine the efficacy of commercial groups and non-commercial groups at 1-year, compared to individual interventions. It was not possible to determine efficacy according to patient-reported outcome measures, including quality of life, as there was insufficient data available (refer to section 3.4).

4.1.1 Commercial Groups versus Individual Interventions

There is a 'high' certainty in the evidence that commercial group interventions are favourable over individual interventions in inducing a clinically meaningful 5% weight loss at 1-year. In context, per 1,000 service users referred for weight management intervention, 217 more service users will achieve 5% weight loss in a commercial group, then in an individual intervention. In other words, according to number needed to treat (NNT) analysis, 1 in 6 service users (95% CI 1 in 4.4 to 8.2) will achieve 5% weight loss by attending a commercial group. The certainty of the evidence was also 'high' to suggest that commercial group interventions are cost-effective when considered against the nominal cost-per-QALY threshold (Appleby et al. 2007).

Superiority in weight loss did not translate into any clinical reduction in risk of cardiovascular disease or diabetes at 1-year, however. There is 'moderate' certainty that neither individual nor commercial group interventions led to any clinical improvement in measures of systolic blood pressure, total:HDL cholesterol or HbA1c.

4.1.2 Non-Commercial Groups versus Individual Interventions

There is ‘moderate’ certainty that indicates a that there was favourability in neither non-commercial groups or individual interventions for the attainment of 5% weight loss; although this is not conclusive. The wide confidence interval suggests that after 1-year, per 1,000 service users, 59 more service users would achieve 5% weight loss in the non-commercial group, compared to individual intervention. However, this risk difference was inconclusive based on a confidence interval whereby anywhere between 14 fewer to 151 more service users would achieve this outcome. NNT was not calculated for non-commercial groups due to the absence of statistical significance.

There was ‘moderate’ certainty in the evidence that there are ambivalent changes to markers of cardiovascular disease and diabetes control, showing no clinical improvements or detriment to total:HDL cholesterol, systolic blood pressure or HbA1c as a result of a non-commercial group compared to individual intervention.

There is ‘high’ certainty that non-commercial group interventions are cost-effective, according to NICE thresholds (Appleby et al. 2007). In fact, commercial and non-commercial group interventions are equally cost-effective.

4.2 Strengths and Limitations

4.2.1 Methodological Rigor

This systematic review followed a protocol that was made publicly available online on the PROSPERO website (Abbott and Bryant 2017), prior to the literature searches being undertaken. Thus, this deters a reporting bias towards significant findings and ensures protocol fidelity. However, the absence of any gender data in the included studies meant that the planned sub-group analysis on gender was omitted. This represents an addendum from

the published protocol. This protocol amendment has been recorded electronically on PROSPERO for audit purposes and is fully transparent and justified.

The search strategy was designed alongside a subject-expert librarian and followed the 'Highly Sensitive Search Strategy' used by the Cochrane Collaboration (Glanville et al. 2006). In addition, to ensure the search was sensitive and thus maximising the yield of relevant studies, comprehensive MeSH and keyword search terms were used. The search strategies were applied to multiple healthcare-orientated databases, thus reducing the risk of introducing a database bias. However, the literature search did rely solely on database searching and did not use other possible sources, such as hand-searching relevant journals; and therefore source selection bias cannot be ruled out.

The clearly defined research question led to detailed criteria for inclusion and exclusion, which were used to accept or reject studies for inclusion as part of the screening process. Screening was conducted by two reviewers, who were blinded to each other's decisions until the point of arbitration. Using at least two reviewers reduces the possibility that relevant literature will be discarded (Edwards et al. 2002) and reduces the risk of selection bias. Further, data extraction was peer reviewed to minimise data errors. The agreement between reviewers was not quantified, however, and it may have been useful to use kappa statistics (Hedges and Cooper 1994) to measure inter-rater reliability.

Further, the reporting on this review followed the evidence-based items for reporting in systematic reviews, as outlined by Moher et al. (2009). This provides transparency and detail to allow the reader to appraise this systematic review and make their own make judgements.

4.2.2 Generalisability

Careful consideration was given to the specification of the inclusion criteria for the review, in order to ensure that the review could be generalised to overweight and obese adults across populations. The inclusion criteria for the intervention setting included any non-inpatient settings and these were not restricted to specific countries.

The population sample within this review included a substantial sample size of 2,139 participants. Interventions predominately took place in the United Kingdom and, as well, were delivered in other westernised countries: United States of America, Spain, Australia and Germany. The setting of intervention was varied across outpatient, primary care and community settings. Moreover, recruitment was pre-dominantly multi-centred and therefore the population within this review is believed to be statistically representative of, and generalisable to, a large westernised population.

Despite inconsistency in the models and programme design of the commercial- and non-commercial groups, results were overall homogenous within each category. The inconsistency in intervention designs is reflective of the diversity in the design of lifestyle interventions internationally. In this, these results are not limited to a specific model of intervention and can be generalisable to commercial or a non-commercial lifestyle group intervention as separate entities; regardless of how they are delivered in practice.

There was no time exclusion for publication within this review; as lifestyle interventions for overweight and obesity are not novel to modern day. All studies were undertaken post-millennial with the exception of Long et al. (1983). It could, however, be argued that since the 1980's the scale and demographic of the overweight and obese population has changed, and therefore these participants may not be representative of today's population.

4.2.3 Applicability

The explicit research question and inclusion asked in this review allows any reader to establish the applicability of this review to their own population group. In relation to practice in the UK, this review is highly applicable.

The demographic of the participants included in this review mirror that of the weight management referral criteria for Tier 2 and Tier 3 services in the UK. Participants within this review were towards middle-age (mean 47 years old), obese (mean BMI 35.5kg/m^2) and the majority were female (67.3%). In context, the demographics of participants of Tier 3 weight management are also the majority female (71%), aged 45-64 years old (45.1%) and have a BMI of 35-40kg/m^2 (35.4%) (Blane et al. 2017). This demographic is echoed by (Ahern et al. 2016) for Tier 2 service users, whereby service users are on average 53 years old, have a BMI of 34.5kg/m^2 and are 67.9% female.

In clinical practice, Tier 2 and Tier 3 services are funded by local authorities and clinical commissioning groups; thus are not centrally funded. The provisions of these services are highly variable and 40% of CCGs and local authorities in England do not commission such services (Coulton et al. 2015). This 'postcode lottery' means that NHS-funded interventions for overweight and obesity are not always accessible to service users, and therefore may not be a viable treatment option. Therefore, although this research suggests that commercial group interventions are the most effective treatment option for weight loss, these may not funded through primary care referral and therefore access may only be possible via the service users' self-funding.

4.2.4 Outcome Measures

International obesity guidelines recommend a weight loss of 5% to achieve improvements in health outcomes (Jensen et al. 2014, NICE 2014a). Thus, this review measured arbitrary achievement of 5% weight loss in order to directly establish clinical efficacy. Only 5 of the 8

studies reported on proportional weight loss in this way. Attempts were made to provide more complete data by conducting a sensitivity analysis using the weight change data reported in kilograms. This sensitivity analysis contained the usable data from 7 of the 8 studies. Measuring absolute, as opposed to proportional, weight loss did not alter the direction, significance or magnitude of effect. This demonstrates that the observed outcomes for achievements in 5% weight loss are robust, regardless of the unit of measurement used.

Many of the clinical outcomes measured in this review were surrogate markers and therefore cannot definitively ascertain diabetes and cardiovascular disease risk. Moreover, even after sub-group and sensitivity analyses were performed, significant heterogeneity remained for blood pressure. This may suggest that there were other factors that could explain the variance; or it could be a reflection of the methodological rigor in which these measurements were obtained.

IPAQ was the sole measure of physical activity used by the included studies in this review. IPAQ is not validated as an outcome measurement for interventional studies; give that there is lack of evidence that IPAQ is responsive to change in physical activities. Therefore, these results may not be valid (Bauman et al. 2009). Considering this, however, this unit of measure is the most feasible and practical to administer in the context of interventional studies; particularly those with larger sample sizes.

While physical activity was uniformly measured across included studies using IPAQ; selective reporting of the IPAQ domains meant that there was no consistency in measure. To enable this data to be used for the meta-analysis, standardised mean difference (SMD) was used as a summary statistic. However, this analysis leads to difficulty in clinical interpretation; as the effect size is measured as SMD rather than in clinical units.

4.2.5 Data Imputations

BOCF data was extracted a priori as it is viewed to be more methodologically robust than completers-only data (Cresswell and Mander 2014). Nevertheless, the validity of this approach for obesity interventions is debatable. In practice, the BOCF approach is highly conservative considering that some participants may regain an excess of the weight lost, after completing their treatment (Ware 2003). Where only completers-only data was available, this was included in the meta-analyses in order to not exclude otherwise eligible studies. It is acknowledged, however, that meta-analysis results may differ according to the missing data assumptions of included studies (Cresswell and Mander, 2014). Sensitivity analysis according to risk of bias addressed this. The Cochrane Risk of Bias Tool includes an assessment of the handling of missing data to determine attrition bias. Studies in this review were judged to have a 'high' risk of bias for the attrition bias domain if missing data imputations were not performed. Therefore, the risk of bias sensitivity analyses within this review provides evidence that combining data from studies with and without imputations did not influence the direction or size of effect.

4.2.6 Publication Bias

While not applying a language restriction to the article inclusion criteria meant screening was required translation and thus added labour, this did reduce language bias. The included articles in this review were all written in the English language which may reflect the *lingua franca* of global science; but this does not mean one should assume that all important research is available in English (Amano et al. 2016). We did not carry out searches of non-English language dominated databases, such as SciELO, and therefore our included studies are biased towards abstracts published in English.

Publication bias was minimised further by searching the ISRCTN database and permitting inclusion of grey literature. Despite this, only one conference proceeding was included as a

linked article to the Jebb et al. (2011) study. Further, some grey literature was excluded as their eligibility was inconclusive and authors did not respond to information requests.

There is indirect evidence of publication bias, given that all but one study was a large study and all reported significant findings. Only observational evidence of publication bias was able to be performed, owing to an insufficient number of studies to test for publication bias through funnel plot asymmetry. Publication bias towards larger studies is not necessarily a limitation, given that larger trials tend to have greater methodological rigor (Egger et al. 2003). Further, the dominance of large studies in this review means that the issue of small study effects and overestimation of effect sizes is minimised.

4.2.7 Intervention Components

This review did not seek to identify which intervention components contributed to efficacy and therefore meta-regression analyses were not performed. Therefore, while this review has firmly established that commercial-group interventions are more effective than individual interventions, it is not clear why.

This review hypothesised that group interventions would be more effective than individual interventions. It is believed that group lifestyle interventions provide peer support which can, in turn, modify the participant's social network (Christakis and Fowler 2007). In this sense, network phenomena and social influence can spread positive weight management and health behaviours between those attending interventions for overweight or obesity.

The findings of this review support this as a possible hypothesis, given that all group interventions included in this review demonstrated the interventionist delivering the 'social support' behavioural taxonomy (Michie et al. 2013) and thus implies the enactment of 'Buddying Support', 'Motivational Support' and 'Imitation Modelling' behaviours; according to OxFAB taxonomy (Hartmann-Boyce et al. 2016).

It is arguable, however, that the treatment effect may not be due to peer support, rather it may be due to intervention intensity. Research has shown that greater weight loss during lifestyle interventions is associated with higher contact time or frequency of contacts per participant (Greaves et al. 2011). As well as providing more frequent contact, group interventions included in this review provided more hours of contact per participant (range of 12 to 55 hours) compared to individual interventions (range of 2.5 to 11 hours). It is therefore a plausible hypothesis that group interventions could be more effective owing to their time-efficiency thereby providing greater interventionist contact time per participant.

4.3 Comparison with Other Systematic Reviews

This systematic review is compared with existing systematic reviews that have directly or in-directly evaluated the efficacy of group-based compared to individually delivered lifestyle interventions for overweight and obese adults.

4.3.1 Direct Comparisons

The systematic review by Paul-Ebhohimhen and Avenell (2009) provided a direct comparison between group and individual interventions. This systematic review paper included citations from two combined published reviews (Avenell et al. 2004, (Paul-Ebhohimhen and Avenell 2008). These citations were obtained using a broad search strategy and included studies up to 2001 (Avenell et al. 2004) and added to by citation alerts for some, but not all databases, to the year 2008. Further, journals were hand-searched specifically for financial reward interventions (Paul-Ebhohimhen and Avenell 2008). The limitations of this broad search strategy meant that there was poor reliability in the results. Reliability was limited by the inclusion of predominately small studies which were dominated by a female population. Moreover, their review lacked recent data; including only one study published after 1986. Applicability of the pooled results to the general overweight and obese population is also limited, given the generally low attrition rates which were as low as 3.4 to 18%. Further, Paul-

Ebhohimhen and Avenell (2009) collected anthropometric data only. The authors had planned to conduct cost-effectiveness analyses; however this was not possible owing to insufficient data. Their lack of outcome data on cost-efficacy is not surprising given that pre-millennial, conducting cost-effectiveness analysis to establish the 'value for money' of medical interventions was not established as an objective method (Kirkdale et al. 2010).

In addition, the authors did not extract data on proportional weight loss. However, Paul-Ebhohimhen and Avenell (2009) did report summary estimates on weight loss, expressed as an absolute weight loss in kilograms. Results from their meta-analysis of 11 intervention groups showed homogenous findings that group interventions led to a significantly greater weight loss (p= 0.03) at one year, when compared with individual treatment (-1.4kg, 95% CI -2.7kg to -0.1kg). Comparing their results with our sensitivity analysis of 7 intervention groups shows that our results were similar. Albeit, our findings favoured group over individual interventions with more significance (p= <0.00001). Moreover, we found participants lost nearly a further kilogram more than participants in the Ebhohimhen and Avenell (2009) study. Our findings also had a greater degree of precision, with a narrower confidence interval (-2.32kg, 95% CI -2.26kg to -2.19kg). However, our results were highly heterogeneous, even when based on low risk of bias studies (I^2 98%).

Therefore, our review provides an up-to date, and arguably more holistic, examination of efficacy beyond anthropometry alone and we provided a cost-effectiveness analysis to inform policy. In comparison, all but one of the studies in our review were large studies, mixed-gender and published post-millennial (2001 to 2016) which is owed to our more robust methodology. Further, the differences in methodologies between our studies explains why, despite having similar research questions and inclusion/exclusion criteria, our review included only one mutual study (Long et al. 1983) and why our included studies were substantially more modern.

4.3.2 Indirect Comparisons

Systematic reviews of RCTs with head-to-head direct comparisons, such as Paul-Ebhohimhen and Avenell (2009) and our own review, reduce the potential influence of confounders that may contribute to differences in effect; by ensuring participants are randomised using the same criteria. However, systematic reviews with indirect comparisons can include a wider pool of evidence and therefore also have their place.

Hartmann-Boyce and colleagues produced two such systematic reviews which provided indirect comparisons between group and individual interventions (Hartmann-Boyce et al. 2014a) and between commercial and primary-care delivered interventions (Hartmann-Boyce et al. 2014b).

The primary outcome for both studies was absolute weight loss, measured in kilograms. The authors acknowledged that the power of their reviews to detect changes in secondary outcomes, such as blood pressure, lipid profile and fasting glucose, was limited due to a lack of data reporting. Our study has provided more power to detect changes in some, but not all, of these secondary outcomes given that our literature search identified more recent studies (up to 2016), compared to Hartmann-Boyce and colleagues reviews whose included studies were up to 2012. Like our review, however, Hartmann-Boyce and colleagues found insufficient evidence on quality of life or mental health. Further, like us, they found no evidence that programmes led to adverse events.

Hartmann-Boyce and colleagues used a consistent approach to outcome data collection using BOCF as a means of imputation for all studies, as described by Kaiser et al. (2012). This provides a conservative estimate suitable for assessing the benefit to the obese population. It was not possible to use this method in our review, as our primary outcome examined proportional, rather than absolute, weight loss. In addition, and likewise to our methods,

Hartmann-Boyce and colleagues conducted sensitivity analysis using only complete-case data; which also showed similar results.

Hartmann-Boyce et al. (2014a) conducted meta-regression analyses to examine the characteristics of lifestyle interventions that determine efficacy in weight loss, measured in kilograms. While we did not perform regression analyses in our review, we can draw associations between the results from Hartmann-Boyce et al. (2014a) and our own. However, it should be reiterated that the results from Hartmann-Boyce et al. (2014a) are based on in-direct comparisons and therefore are not an exact comparison to our own study.

The regression analyses performed were of the delivery method, as categorised by group or individual. Sixteen interventions were delivered by group sessions while 21 were delivered via individual sessions. Neither group (0.4kg, 95% CI -1.6 to 2.3kg, p=0.71) nor individually (-0.4kg, 95% CI -1.9 to 2.0kg, p=0.97) delivered interventions were found to be significant contributors to weight loss efficacy. These results differ from our findings which, based on a direct comparison, found that group interventions were significantly favoured over individual interventions (p= <0.00001); albeit based upon heterogeneous results. Heterogeneity in our study was then explained by provider sub-group analysis. Hartmann-Boyce et al. (2014a) did not analyse commercial and non-commercial groups as separate entities and therefore comparisons cannot be made in this regard.

Regression analyses were also performed according to behavioural taxonomy, as categorised by CALOR-E (Michie et al. 2013); finding that only the 'comparison of behaviour' was associated with greater weight loss in kilograms (-1.5kg, 95% CI -2.9 to -0.1kg, p= 0.032). Meanwhile, the 'social support' taxonomy was not associated with weight change (0.5kg, 95% CI -0.6 to 1.6kg, p= 0.360). Results from their regression analysis do not support our hypothesis that the 'social support' taxonomy could explain the greater efficacy of group interventions. However, 5 of the 12 group interventions included in our review were also

identified to use the ‘comparison on behaviour’ taxonomy. Interestingly, this taxonomy is a consistent component of the ‘WeightWatchers’ commercial group programme. Thus, this raises an alternative underpinning that participants are motivated by comparing their behaviour, or weight loss results, with their peers; driving further weight loss or maintenance.

It should be reiterated that these comparisons should be interpreted with caution. In-direct comparisons lose the power of randomisation and are likely to be biased (Bucher et al. 1997).

A further study, by Hartmann-Boyce et al. (2014b), performed meta-analyses sub-grouped by intervention provider. They categorised interventions as commercial (group-delivered) or primary-care (individual- or group-delivered). The results from the primary-care sub-group analyses were not exclusively individual interventions and therefore are irrelevant to our review and will not be discussed. The commercial sub-group, on the other hand, is relevant.

Their commercial sub-group included 5 commercial group programmes. Results from their meta-analysis showed that commercial interventions are effective compared to minimal intervention; yielding a mean difference of -2.22kg (95% CI -2.89kg to -1.54kg). All included studies were common with the commercial studies included in our review (Heshka et al. 2003; Jolly et al. 2011; Jebb et al. 2011). Our study compared against an active individual intervention. We found that commercial group interventions were equally significant (p= <0.00001). Findings from our sensitivity analysis using absolute weight loss, found a similar mean weight loss of -2.29kg (95% CI -2.33 to 2.25kg), but this was not identical. Further, our results were entirely homogenous (I^2 0%) while heterogeneity remained (I^2 20%) in findings by Hartmann-Boyce et al. (2014b). These discrepancies are because of our systematic review directly comparing with an active individual intervention rather than against a minimal intervention control; as Hartmann-Boyce and colleagues did.

5.0 Conclusions

This is the first systematic review for nearly a decade to directly determine the effectiveness of group compared to individual lifestyle interventions. This research has added new knowledge, determining that commercialised group programmes have a profound and beneficial influence on weight loss efficacy. The implications for practice have been considered below in context for service users, clinicians and policy makers in the UK; given that this was the dominant setting for the studies included within this review. However, it should be noted that this review is also generalisable to other westernised populations.

5.1 For Service users

Adult service users who are overweight or obese should speak with their primary care provider to establish which lifestyle interventions are available to them, as there is substantial geographical variation in access. If a service user has access to several options, the findings of this review suggest that service users should choose a commercial group over a non-commercial group or individual intervention.

Service users should not be discouraged, however, if commercial group interventions are not available to them through the NHS. This review found that non-commercial group and individual interventions also led to weight loss, albeit less significant. What is more, group interventions may not be suited to all service users; including service users suffering from agoraphobia or social anxiety, or service users requiring translator services.

Service users tend to place priority on quality of life outcomes, particularly outcomes related to emotional and psychological functioning, when considering the efficacy of treatment options (Janse et al. 2004). Whilst we pre-specified quality of life and mental health as secondary outcomes in this review, there was insufficient evidence available to determine effectiveness as measured by patient-reported outcome measures.

While the interventions included in this review induced weight loss, no other clinical benefit was seen. Therefore, service users should be realistic about the health improvements that they will see from a lifestyle intervention alone over a 1-year period. Considering this, service users with co-morbidities should be open to pharmacological and/or surgical interventions as concomitant treatment alongside lifestyle intervention.

5.2 For Clinicians

Current NICE guidance (NICE 2014a) are limited in their interpretation for lifestyle interventions. The knowledge that this present systematic review contributes will compliment NICE guidance by providing more detailed and practical information to help clinicians consider the most effective modality of lifestyle intervention for their service users.

The evidence from this review suggests that clinicians should favour referral to a group rather than individual lifestyle intervention. More specifically, this evidence suggests that clinicians in primary care should consider referral to commercial over non-commercial group intervention; if the option is available. This supports previous research that has shown that referral to commercial weight management services from primary care are suitable, practical and effective (Lewis et al. 2013; Aveyard et al. 2016).

While the evidence from this review concludes that commercial group interventions are the most effective treatment modality for the general population, the individual service user presenting in clinical practice may not feel this option is suited to them. Treatment should be a shared decision-making process and service users' choices should be exercised to promote treatment fidelity.

5.3 For Policy Makers

Even though this review concluded with high certainty that commercial group interventions are shown to be more effective in attaining a clinically important 5% weight loss, policy makers will ultimately make their decisions based on cost-effectiveness. The evidence from this review has demonstrated that both commercial and non-commercial group interventions are cost-effective by cost per QALY, according to NICE thresholds (Appleby et al. 2007).

Commissioners are likely to prioritise cost per QALY gained in order to maximise health gain for the whole population. Given that obesity is a public health epidemic, the commissioning priority may be to reduce health inequalities across the population. In this case, commissioners may value a cost-consequence analysis to base their decision-making (NICE 2011b); which unfortunately is not offered in addition within this review.

6.0 Future Research

The implication of this systematic review in guiding future research has highlighted the need for weight management interventions to report to a core outcome set. While at present a core outcome set does not exist, partly because of the complexity and variability of intervention targets, national guidance frameworks do exist on minimal outcome reporting i.e. the Standard Evaluation Framework for Weight Management Interventions (Ells et al. 2018; updated from Roberts et al. 2012). In order to enhance the comparability and synthesis of future studies, future empirical studies should follow guidance on the minimal outcome reporting recommendations.

Given that obesity is a chronic condition and weight maintenance after initial weight loss is a challenge, RCTs with prolonged follow-up are required to establish efficacy in the longer-term. Future RCTs should also consider reporting proportional weight loss, such as achievement of 5% weight loss, in order to provide more clinically relevant data. Including economic analyses in future studies will increase the evidence base on cost-effectiveness and will be necessary to enable clinicians and policy makers to manage service provision and make the best use of resources. Further, there is a need for greater emphasis on patient-reported outcome measures to assess efficacy based on service users' concerns; rather than focusing on the clinical outcomes believed to be prudent to researchers.

This systematic review did not directly compare the efficacy of commercial and non-commercial group interventions with each other. A future RCT comparing these two intervention providers, with equal time and contact intensity, would be the means to find a definitive answer as to whether commercial group-interventions are truly more effective than non-commercial group interventions.

Moreover, further research could also seek to identify interventional components that determine commercial group interventions to be more effective. This knowledge would enable replication in non-commercial led group interventions.

On the other hand, this review has established with ‘high’ certainty that commercial group interventions are superior for weight loss, as opposed to individual interventions. This review has also provided in-direct evidence that commercial group interventions are equally as cost-effective as, and more widely accessible than, non-commercial group interventions.

Therefore, I would raise the question as to whether the cost of further research is even warranted or justified. Rather, this review may provide sufficient evidence to convince policy makers that the first-line lifestyle intervention for adults with overweight and obesity should be a commercial group programme.

7.0 References

Abbott, S. and Bryant, E. (2017) *PROSPERO: International Prospective Register of Systematic Reviews. CRD42017056396.* [online] available from <https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=56396>

Adrian Bauman, Barbara E. Ainsworth, Fiona Bull, Cora L. Craig, M.H. and James F. Sallis, Michael Pratt, and M.S. (2009) 'Progress and Pitfalls in the Use of the International Physical Activity Questionnaire (IPAQ) for Adult Physical Activity Surveillance'. *Journal of Physical Activity and Health* [online] 6, 5–8. available from <https://s3.amazonaws.com/academia.edu.documents/41508616/Physical_Activity_1_Global_physical_acti20160124-11229-iud5a9.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1523384512&Signature=Ig7f4t6%2F6OXk2NHicpGAaqzijIc%3D&response-content-disposition=inlin> [10 April 2018]

Ahern, A.L., Aveyard, P., Boyland, E.J., Halford, J.C., Jebb, S.A., and WRAP trial team (2016) 'Inequalities in the Uptake of Weight Management Interventions in a Pragmatic Trial: An Observational Study in Primary Care.' *The British Journal of General Practice : The Journal of the Royal College of General Practitioners* [online] 66 (645), e258-63. available from <http://www.ncbi.nlm.nih.gov/pubmed/26906629> [8 March 2018]

Ali, M.K., Echouffo-Tcheugui, J.B., and Williamson, D.F. (2012) 'How Effective Were Lifestyle Interventions In Real-World Settings That Were Modeled On The Diabetes Prevention Program?' *Health Affairs* [online] 31 (1), 67–75. available from <http://www.ncbi.nlm.nih.gov/pubmed/22232096> [25 April 2018]

Amano, T., González-Varo, J.P., and Sutherland, W.J. (2016) 'Languages Are Still a Major Barrier to Global Science'. *PLOS Biology* [online] 14 (12), e2000933. available from

<http://dx.plos.org/10.1371/journal.pbio.2000933> [16 March 2018]

Appel, L.J., Clark, J.M., Yeh, H.-C., Wang, N.-Y., Daumit, G., Miller III, E.R., Brancati, F.L., Dalcin, A., Jerome, G.J., Noronha, G., Pozefsky, T., Reynolds, J.B., Durkin, N., Coughlin, J.W., Rubin, R.R., Charleston, J., Louis, T.A., and Geller, S. (2011) 'Comparative Effectiveness of Weight-Loss Interventions in Clinical Practice'. *New England Journal of Medicine* 365 (21), 1959–1968

Appleby, J., Devlin, N., and Parkin, D. (2007) 'NICE's Cost Effectiveness Threshold.' *BMJ (Clinical Research Ed.)* [online] 335 (7616), 358–9. available from <http://www.ncbi.nlm.nih.gov/pubmed/17717337> [23 March 2018]

Ash, S., Reeves, M., Bauer, J., Dover, T., Vivanti, A., Leong, C., Sullivan, T.O., and Capra, S. (2006) 'A Randomised Control Trial Comparing Lifestyle Groups, Individual Counselling and Written Information in the Management of Weight and Health Outcomes over 12 Months'. *International Journal of Obesity* [online] 30 (10), 1557–1564. available from <http://www.ncbi.nlm.nih.gov/pubmed/16534529> [23 May 2017]

Bandura, A. (1997) *Self-Efficacy : The Exercise of Control*.

Blane, D.N., McLoone, P., Morrison, D., Macdonald, S., and O'Donnell, C.A. (2017) 'Patient and Practice Characteristics Predicting Attendance and Completion at a Specialist Weight Management Service in the UK: A Cross-Sectional Study.' *BMJ Open* [online] 7 (11), e018286. available from <http://www.ncbi.nlm.nih.gov/pubmed/29162575> [8 March 2018]

Booth, H.P., Prevost, T.A., Wright, A.J., and Gulliford, M.C. (2014) 'Effectiveness of Behavioural Weight Loss Interventions Delivered in a Primary Care Setting: A Systematic Review and Meta-Analysis'. *Family Practice* [online] 31 (6), 643–653.

available from <https://academic.oup.com/fampra/article-lookup/doi/10.1093/fampra/cmu064> [23 April 2018]

British Cardiac Society, B.H.S.D.U.H.U.P.C.C.S.T.S.A., British Hypertension Society, Diabetes UK, HEART UK, Primary Care Cardiovascular Society, and Stroke Association (2005) 'JBS 2: Joint British Societies' Guidelines on Prevention of Cardiovascular Disease in Clinical Practice.' *Heart (British Cardiac Society)* [online] 91 Suppl 5 (suppl 5), v1-52. available from <http://www.ncbi.nlm.nih.gov/pubmed/16365341> [31 January 2018]

Bucher, H.C., Guyatt, G.H., Griffith, L.E., and Walter, S.D. (1997) 'The Results of Direct and Indirect Treatment Comparisons in Meta-Analysis of Randomized Controlled Trials'. *Journal of Clinical Epidemiology* [online] 50 (6), 683–691. available from <https://www.sciencedirect.com/science/article/pii/S0895435697000498> [12 April 2018]

Butryn, M.L., Webb, V., and Wadden, T.A. (2011) 'Behavioral Treatment of Obesity'. *Psychiatric Clinics of North America* [online] 34 (4), 841–859. available from <http://www.ncbi.nlm.nih.gov/pubmed/22098808> [25 April 2018]

Centre for Reviews and Dissemination. (2009) *Systematic Reviews : CRD's Guidance for Undertaking Reviews in Health Care*. CRD, University of York

Christakis, N.A. and Fowler, J.H. (2007) 'The Spread of Obesity in a Large Social Network over 32 Years'. *New England Journal of Medicine* [online] 357 (4), 370–379. available from <http://www.ncbi.nlm.nih.gov/pubmed/17652652> [23 May 2017]

Cochrane Effective Practice and Organisation of Care (2017) *EPOC Resources for Review Authors* [online] available from <http://epoc.cochrane.org/epoc-specific-resources-

review-authors>

Coulton, V., Dodhia, S., Ells, L., and Blackshaw, J. (2015) *National Mapping of Weight Management Services Provision of Tier 2 and Tier 3 Services in England*. [online] available from <www.gov.uk/phe> [25 March 2018]

Cresswell, L. and Mander, A.P. (2014) 'How to Use Published Complete Case Results from Weight Loss Studies in a Missing Data Sensitivity Analysis'. *Obesity* [online] 22, 996–1001. available from <https://onlinelibrary.wiley.com/doi/pdf/10.1002/oby.20635> [17 April 2018]

Deci, E.L. and Ryan, R.M. (2000) 'The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior'. *Psychological Inquiry* [online] 11 (4), 227–268. available from <http://www.tandfonline.com/doi/abs/10.1207/S15327965PLI1104_01> [25 April 2018]

Deeks, J.J. (2002) 'Issues in the Selection of a Summary Statistic for Meta-Analysis of Clinical Trials with Binary Outcomes'. *Statistics in Medicine* [online] 21 (11), 1575–1600. available from <http://doi.wiley.com/10.1002/sim.1188> [17 April 2018]

DerSimonian, R. and Laird, N. (1986) 'Meta-Analysis in Clinical Trials.' *Controlled Clinical Trials* [online] 7 (3), 177–88. available from <http://www.ncbi.nlm.nih.gov/pubmed/3802833> [17 April 2018]

Dossing, A., Tarp, S., Furst, D.E., Gluud, C., Beyene, J., Hansen, B.B., Bliddal, H., and Christensen, R. (2014) 'Interpreting Trial Results Following Use of Different Intention-to-Treat Approaches for Preventing Attrition Bias: A Meta-Epidemiological Study Protocol.' *BMJ Open* [online] 4 (9), e005297. available from <http://www.ncbi.nlm.nih.gov/pubmed/25260368> [17 April 2018]

Easterbrook, P.J., Berlin, J.A., Gopalan, R., and Matthews, D.R. (1991) 'Publication Bias in Clinical Research.' *Lancet (London, England)* [online] 337 (8746), 867–72. available from <http://www.ncbi.nlm.nih.gov/pubmed/1672966> [7 November 2017]

Edwards, P., Clarke, M., DiGuiseppi, C., Pratap, S., Roberts, I., and Wentz, R. (2002) 'Identification of Randomized Controlled Trials in Systematic Reviews: Accuracy and Reliability of Screening Records'. *Statistics in Medicine* [online] 21 (11), 1635–1640. available from <http://www.ncbi.nlm.nih.gov/pubmed/12111924> [10 April 2018]

Egger, M., Juni, P., Bartlett, C., Holenstein, F., and Sterne, J. (2003) 'How Important Are Comprehensive Literature Searches and the Assessment of Trial Quality in Systematic Reviews? Empirical Study.' *Health Technology Assessment (Winchester, England)* [online] 7 (1), 1–76. available from <http://www.ncbi.nlm.nih.gov/pubmed/12583822> [10 April 2018]

Elfhag, K. and Rossner, S. (2005) 'Who Succeeds in Maintaining Weight Loss? A Conceptual Review of Factors Associated with Weight Loss Maintenance and Weight Regain'. *Obesity Reviews* [online] 6 (1), 67–85. available from <http://www.ncbi.nlm.nih.gov/pubmed/15655039> [25 April 2018]

Elobeid, M.A., Padilla, M.A., McVie, T., Thomas, O., Brock, D.W., Musser, B., Lu, K., Coffey, C.S., Desmond, R.A., St-Onge, M.-P., Gadde, K.M., Heymsfield, S.B., and Allison, D.B. (2009) 'Missing Data in Randomized Clinical Trials for Weight Loss: Scope of the Problem, State of the Field, and Performance of Statistical Methods'. *PLoS ONE* [online] 4 (8), e6624. available from <http://dx.plos.org/10.1371/journal.pone.0006624> [17 April 2018]

Finley, C.E., Barlow, C.E., Greenway, F.L., Rock, C.L., Rolls, B.J., and Blair, S.N. (2007) 'Retention Rates and Weight Loss in a Commercial Weight Loss Program'. *International*

Journal of Obesity [online] 31 (2), 292–298. available from <http://www.nature.com/articles/0803395> [25 April 2018]

Franz, M.J., VanWormer, J.J., Crain, A.L., Boucher, J.L., Histon, T., Caplan, W., Bowman, J.D., and Pronk, N.P. (2007) 'Weight-Loss Outcomes: A Systematic Review and Meta-Analysis of Weight-Loss Clinical Trials with a Minimum 1-Year Follow-Up'. *Journal of the American Dietetic Association* [online] 107 (10), 1755–1767. available from <http://www.ncbi.nlm.nih.gov/pubmed/17904936> [1 August 2017]

Galani, C. and Schneider, H. (2007) 'Prevention and Treatment of Obesity with Lifestyle Interventions: Review and Meta-Analysis.' *International Journal of Public Health* [online] 52 (6), 348–59. available from <http://www.ncbi.nlm.nih.gov/pubmed/18368998> [23 April 2018]

Glanville, J.M., Lefebvre, C., Miles, J.N. V, and Camosso-Stefinovic, J. (2006) 'How to Identify Randomized Controlled Trials in MEDLINE: Ten Years On.' *Journal of the Medical Library Association : JMLA* [online] 94 (2), 130–6. available from <http://www.ncbi.nlm.nih.gov/pubmed/16636704> [10 April 2018]

Gortmaker, S.L., Swinburn, B.A., Levy, D., Carter, R., Mabry, P.L., Finegood, D.T., Huang, T., Marsh, T., and Moodie, M.L. (2011) 'Changing the Future of Obesity: Science, Policy, and Action.' *Lancet (London, England)* [online] 378 (9793), 838–47. available from <http://www.ncbi.nlm.nih.gov/pubmed/21872752> [25 April 2018]

Great Britain. Department of Health. and Panel on Dietary reference Values of the Committee on Medical Aspects of Food Policy. (1991) *Dietary Reference Values for Food Energy and Nutrients for the United Kingdom : Report of the Panel on Dietary Reference Values of the Committee on Medical Aspects of Food Policy*. H.M.S.O

Greaves, C.J., Sheppard, K.E., Abraham, C., Hardeman, W., Roden, M., Evans, P.H., and Schwarz, P. (2011) 'Systematic Review of Reviews of Intervention Components Associated with Increased Effectiveness in Dietary and Physical Activity Interventions'. *BMC Public Health* [online] 11 (1), 119. available from <http://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-11-119> [8 April 2018]

Grover, S.A., Kaouache, M., Rempel, P., Joseph, L., Dawes, M., Lau, D.C.W., and Lowensteyn, I. (2015) 'Years of Life Lost and Healthy Life-Years Lost from Diabetes and Cardiovascular Disease in Overweight and Obese People: A Modelling Study'. *The Lancet Diabetes & Endocrinology* [online] 3 (2), 114–122. available from <http://www.ncbi.nlm.nih.gov/pubmed/25483220> [2 May 2018]

Hartmann-Boyce, J., Aveyard, P., Koshiaris, C., and Jebb, S.A. (2016) 'Development of Tools to Study Personal Weight Control Strategies: OxFAB Taxonomy.' *Obesity (Silver Spring, Md.)* [online] 24 (2), 314–20. available from <http://www.ncbi.nlm.nih.gov/pubmed/26748902> [12 April 2018]

Hartmann-Boyce, J., Johns, D.J., Jebb, S.A., Aveyard, P., and Behavioural Weight Management Review Group (2014a) 'Effect of Behavioural Techniques and Delivery Mode on Effectiveness of Weight Management: Systematic Review, Meta-Analysis and Meta-Regression'. *Obesity Reviews* [online] 15 (7), 598–609. available from <http://www.ncbi.nlm.nih.gov/pubmed/24636238> [6 April 2018]

Hartmann-Boyce, J., Johns, D.J., Jebb, S.A., Summerbell, C., Aveyard, P., and Behavioural Weight Management Review Group (2014b) 'Behavioural Weight Management Programmes for Adults Assessed by Trials Conducted in Everyday Contexts: Systematic Review and Meta-Analysis'. *Obesity Reviews* [online] 15 (11), 920–932. available from <http://www.ncbi.nlm.nih.gov/pubmed/25112559> [6 April 2018]

Hassan, Y., Head, V., Jacob, D., Bachmann, M.O., Diu, S., and Ford, J. (2016) 'Lifestyle Interventions for Weight Loss in Adults with Severe Obesity: A Systematic Review'. *Clinical Obesity* [online] 6 (6), 395–403. available from <http://www.ncbi.nlm.nih.gov/pubmed/27788558> [6 April 2018]

Hayaki, J. and Brownell, K.D. (1996) 'Behaviour Change in Practice: Group Approaches.' *International Journal of Obesity and Related Metabolic Disorders : Journal of the International Association for the Study of Obesity* [online] 20 Suppl 1, S27-30. available from <http://www.ncbi.nlm.nih.gov/pubmed/8646262> [25 April 2018]

Hedges, L. V. and Cooper, H.M. (1994) *The Handbook of Research Synthesis* [online] Russell Sage Foundation. available from <https://books.google.co.uk/books?hl=en&lr=&id=p-aFAwAAQBAJ&oi=fnd&pg=PA125&dq=valuing+coding+decisions+Handbook+of+Research+Synthesis+orwin+1994&ots=TVwMPicLTL&sig=NK1OXAznfXQMtKaNEQo-q_NtAVw#v=onepage&q&f=false> [10 April 2018]

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G., and Badia, X. (2011) 'Development and Preliminary Testing of the New Five-Level Version of EQ-5D (EQ-5D-5L)'. *Quality of Life Research* [online] 20 (10), 1727–1736. available from <http://www.ncbi.nlm.nih.gov/pubmed/21479777> [31 January 2018]

Heshka, S., Anderson, J.W., Atkinson, R.L., Greenway, F.L., Hill, J.O., Phinney, S.D., Kolotkin, R.L., Miller-Kovach, K., and Pi-Sunyer, F.X. (2003) 'Weight Loss With Self-Help Compared With a Structured Commercial Program'. *JAMA* [online] 289 (14), 1792. available from <http://www.ncbi.nlm.nih.gov/pubmed/12684357> [30 May 2017]

Higgins, J. and Green, S. (2011) *Cochrane Handbook for Systematic Reviews of Interventions* [online] available from <http://www.handbook.cochrane.org.> [8

November 2016]

Higgins, J.P.T., Altman, D.G., Gøtzsche, P.C., Jüni, P., Moher, D., Oxman, A.D., Savovic, J., Schulz, K.F., Weeks, L., Sterne, J.A.C., Cochrane Bias Methods Group, and Cochrane Statistical Methods Group (2011) 'The Cochrane Collaboration's Tool for Assessing Risk of Bias in Randomised Trials.' *BMJ (Clinical Research Ed.)* [online] 343, d5928. available from <http://www.ncbi.nlm.nih.gov/pubmed/22008217> [16 April 2018]

Janse, A.J., Gemke, R.J.B.J., Uiterwaal, C.S.P.M., van der Tweel, I., Kimpen, J.L.L., and Sinnema, G. (2004) 'Quality of Life: Patients and Doctors Don't Always Agree: A Meta-Analysis'. *Journal of Clinical Epidemiology* [online] 57 (7), 653–661. available from <http://www.ncbi.nlm.nih.gov/pubmed/15358393> [8 April 2018]

Jebb, S.A., Ahern, A.L., Olson, A.D., Aston, L.M., Holzapfel, C., Stoll, J., Amann-Gassner, U., Simpson, A.E., Fuller, N.R., Pearson, S., Lau, N.S., Mander, A.P., Hauner, H., and Caterson, I.D. (2011) 'Primary Care Referral to a Commercial Provider for Weight Loss Treatment versus Standard Care: A Randomised Controlled Trial'. *The Lancet* [online] 378 (9801), 1485–1492. available from <http://www.ncbi.nlm.nih.gov/pubmed/21906798> [30 May 2017]

Jensen, M.D., Ryan, D.H., Apovian, C.M., Ard, J.D., Comuzzie, A.G., Donato, K.A., Hu, F.B., Hubbard, V.S., Jakicic, J.M., Kushner, R.F., Loria, C.M., Millen, B.E., Nonas, C.A., Pi-Sunyer, F.X., Stevens, J., Stevens, V.J., Wadden, T.A., Wolfe, B.M., Yanovski, S.Z., Jordan, H.S., Kendall, K.A., Lux, L.J., Mentor-Marcel, R., Morgan, L.C., Trisolini, M.G., Wnek, J., Anderson, J.L., Halperin, J.L., Albert, N.M., Bozkurt, B., Brindis, R.G., Curtis, L.H., DeMets, D., Hochman, J.S., Kovacs, R.J., Ohman, E.M., Pressler, S.J., Sellke, F.W., Shen, W.-K., Smith, S.C., Tomaselli, G.F., American College of Cardiology/American Heart Association Task Force on Practice Guidelines, W.-K., Obesity Society, S.O.P., Smith, S.C., Jr, and Tomaselli, G.F. (2014) '2013

AHA/ACC/TOS Guideline for the Management of Overweight and Obesity in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society.' *Circulation* [online] 129 (25 Suppl 2), S102-38. available from <http://www.ncbi.nlm.nih.gov/pubmed/24222017> [27 March 2018]

Jia, H. and Lubetkin, E.I. (2005) 'The Impact of Obesity on Health-Related Quality-of-Life in the General Adult US Population'. *Journal of Public Health* [online] 27 (2), 156–164. available from <http://www.ncbi.nlm.nih.gov/pubmed/15820993> [25 April 2018]

Johns, D.J., Hartmann-Boyce, J., Jebb, S.A., Aveyard, P., and Behavioural Weight Management Review Group (2014) 'Diet or Exercise Interventions vs Combined Behavioral Weight Management Programs: A Systematic Review and Meta-Analysis of Direct Comparisons'. *Journal of the Academy of Nutrition and Dietetics* [online] 114 (10), 1557–1568. available from <http://www.ncbi.nlm.nih.gov/pubmed/25257365> [6 April 2018]

Jolly, K., Lewis, A., Beach, J., Denley, J., Adab, P., Deeks, J.J., Daley, A., and Aveyard, P. (2011) 'Comparison of Range of Commercial or Primary Care Led Weight Reduction Programmes with Minimal Intervention Control for Weight Loss in Obesity: Lighten Up Randomised Controlled Trial'. *BMJ : British Medical Journal (Online)* [online] 343. available from <http://search.proquest.com/docview/1778063255/>

Kaiser, K.A., Affuso, O., Beasley, T.M., and Allison, D.B. (2012) 'Getting Carried Away: A Note Showing Baseline Observation Carried Forward (BOCF) Results Can Be Calculated from Published Complete-Cases Results'. *International Journal of Obesity* [online] 36 (6), 886–889. available from <http://www.ncbi.nlm.nih.gov/pubmed/21407169> [2 May 2018]

Kelly, T., Yang, W., Chen, C.-S., Reynolds, K., and He, J. (2008) 'Global Burden of Obesity in 2005 and Projections to 2030'. *International Journal of Obesity* [online] 32 (9), 1431–1437. available from <http://www.nature.com/doifinder/10.1038/ijo.2008.102> [23 April 2018]

Kirk, S.F.L., Penney, T.L., McHugh, T.-L., and Sharma, A.M. (2012) 'Effective Weight Management Practice: A Review of the Lifestyle Intervention Evidence'. *International Journal of Obesity* [online] 36 (2), 178–185. available from <http://www.ncbi.nlm.nih.gov/pubmed/21487396> [25 April 2018]

Kirkdale, R., Krell, J., O'Hanlon Brown, C., Tuthill, M., and Waxman, J. (2010) 'The Cost of a QALY'. *QJM* [online] 103 (9), 715–720. available from <https://academic.oup.com/qjmed/article-lookup/doi/10.1093/qjmed/hcq081> [12 April 2018]

Kirkham, J.J., Dwan, K.M., Altman, D.G., Gamble, C., Dodd, S., Smyth, R., and Williamson, P.R. (2010) 'The Impact of Outcome Reporting Bias in Randomised Controlled Trials on a Cohort of Systematic Reviews.' *BMJ (Clinical Research Ed.)* [online] 340, c365. available from <http://www.ncbi.nlm.nih.gov/pubmed/20156912> [16 April 2018]

Kroenke, K., Strine, T.W., Spitzer, R.L., Williams, J.B.W., Berry, J.T., and Mokdad, A.H. (2009) 'The PHQ-8 as a Measure of Current Depression in the General Population.' *Journal of Affective Disorders* [online] 114 (1–3), 163–73. available from <http://www.ncbi.nlm.nih.gov/pubmed/18752852> [31 January 2018]

Latner, J.D., Stunkard, A.J., Wilson, G.T., and Jackson, M.L. (2006) 'The Perceived Effectiveness of Continuing Care and Group Support in the Long-Term Self-Help Treatment of Obesity*'. *Obesity* [online] 14 (3), 464–471. available from <http://www.ncbi.nlm.nih.gov/pubmed/16648618> [25 April 2018]

Leahey, T.M. and Wing, R.R. (2013) ‘A Randomized Controlled Pilot Study Testing Three Types of Health Coaches for Obesity Treatment: Professional, Peer, and Mentor’. *Obesity* [online] 21 (5), 928–934. available from <http://doi.wiley.com/10.1002/oby.20271> [25 April 2018]

Long, C.G., Simpson, C.M., and Allott, E.A. (1983) ‘Psychological and Dietetic Counselling Combined in the Treatment of Obesity: A Comparative Study in a Hospital Outpatient Clinic.’ *Human Nutrition. Applied Nutrition* [online] 37 (2), 94–102. available from <http://www.ncbi.nlm.nih.gov/pubmed/6408031> [23 May 2017]

Mantel, N. and Haenszel, W. (1959) ‘Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease’. *JNCI: Journal of the National Cancer Institute* [online] 22 (4), 719–748. available from <https://academic.oup.com/jnci/article/22/4/719/900746/Statistical-Aspects-of-the-Analysis-of-Data-From> [17 April 2018]

Markus, H. and Nurius, P. (1986) ‘Possible Selves.’ *American Psychologist* [online] 41 (9), 954–969. available from <http://doi.apa.org/getdoi.cfm?doi=10.1037/0003-066X.41.9.954> [25 April 2018]

McKinsey Global Institute (2014) *Overcoming Obesity: An Initial Economic Analysis* [online] available from <https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/how-the-world-could-better-fight-obesity> [25 April 2018]

McRobbie, H., Hajek, P., Peerbux, S., Kahan, B.C., Eldridge, S., Trépel, D., Parrott, S., Griffiths, C., Snuggs, S., and Myers Smith, K. (2016) ‘Tackling Obesity in Areas of High Social Deprivation: Clinical Effectiveness and Cost-Effectiveness of a Task-Based Weight Management Group Programme – a Randomised Controlled Trial and Economic Evaluation’. *Health Technology Assessment* [online] 20 (79), 1–150.

available from <http://www.ncbi.nlm.nih.gov/pubmed/27802843> [7 November 2017]

Melmed, S. and Williams, R.H. (2011) *Williams Textbook of Endocrinology.* Elsevier/Saunders

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M.P., Cane, J., and Wood, C.E. (2013) 'The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions'. *Annals of Behavioral Medicine* [online] 46 (1), 81–95. available from <http://www.ncbi.nlm.nih.gov/pubmed/23512568> [22 November 2017]

Moher D, Liberati A, Tetzlaff J, A.D. (2009) 'PRISMA 2009 Flow Diagram'. in *The PRISMA Statement*. vol. 6. 1000097

National Obesity Observatory (2011) *Obesity and Mental Health* [online] Oxford. available from <https://www.scie-socialcareonline.org.uk/obesity-and-mental-health/r/a11G00000017trJIAQ>

Newcombe, R.G. and Bender, R. (2014) 'Implementing GRADE: Calculating the Risk Difference from the Baseline Risk and the Relative Risk.' *Evidence-Based Medicine* [online] 19 (1), 6–8. available from <http://www.ncbi.nlm.nih.gov/pubmed/23970740> [17 April 2018]

Ng, J.Y.Y., Ntoumanis, N., and Thøgersen-Ntoumani, C. (2014) 'Autonomy Support and Control in Weight Management: What Important Others Do and Say Matters'. *British Journal of Health Psychology* [online] 19 (3), 540–552. available from <http://www.ncbi.nlm.nih.gov/pubmed/23751140> [25 April 2018]

NICE (2016) *Cardiovascular Disease: Risk Assessment and Reduction, Including Lipid*

Modification | *Guidance and Guidelines* | *NICE*. [online] available from <https://www.nice.org.uk/guidance/cg181/chapter/1-recommendations> [18 April 2018]

NICE (2015) *Type 2 Diabetes in Adults: Management* | *Guidance and Guidelines* | *NICE*. [online] available from <https://www.nice.org.uk/guidance/ng28> [31 January 2018]

NICE (2014a) *Obesity: Identification, Assessment and Management* | *Guidance and Guidelines* | *NICE*. [online] available from <https://www.nice.org.uk/guidance/cg189/chapter/2-research-recommendations>

NICE (2014b) *Weight Management: Lifestyle Services for Overweight or Obese Adults* | *Guidance and Guidelines* | *NICE*. [online] available from <https://www.nice.org.uk/guidance/ph53/chapter/1-Recommendations> [27 March 2018]

NICE (2013) *Preventing Obesity and Helping People to Manage Their Weight*. [online] available from <https://www.nice.org.uk/advice/lgb9> [25 April 2018]

NICE (2012) *Workplace Health*. [online] available from <https://www.nice.org.uk/advice/lgb2> [25 April 2018]

NICE (2011) *Hypertension in Adults: Diagnosis and Management* | *Guidance and Guidelines* | *NICE*. [online] available from <https://www.nice.org.uk/guidance/CG127/chapter/1-Guidance#diagnosing-hypertension-2> [31 January 2018]

Palmeira, A.L., Teixeira, P.J., Branco, T.L., Martins, S.S., Minderico, C.S., Barata, J.T., Serpa, S.O., and Sardinha, L.B. (2007) 'Predicting Short-Term Weight Loss Using Four Leading Health Behavior Change Theories.' *The International Journal of Behavioral Nutrition and Physical Activity* [online] 4, 14. available from <http://www.ncbi.nlm.nih.gov/pubmed/17448248> [25 April 2018]

Paul-Ebhohimhen, V. and Avenell, A. (2009) ‘A Systematic Review of the Effectiveness of Group versus Individual Treatments for Adult Obesity.’ *Obesity Facts* [online] 2 (1), 17–24. available from <http://www.ncbi.nlm.nih.gov/pubmed/20054200> [7 November 2017]

Paul-Ebhohimhen, V. and Avenell, A. (2008) ‘Systematic Review of the Use of Financial Incentives in Treatments for Obesity and Overweight’. *Obesity Reviews* [online] 9 (4), 355–367. available from <http://www.ncbi.nlm.nih.gov/pubmed/17956546> [23 May 2017]

Peirson, L., Douketis, J., Ciliska, D., Fitzpatrick-Lewis, D., Ali, M.U., and Raina, P. (2014) ‘Treatment for Overweight and Obesity in Adult Populations: A Systematic Review and Meta-Analysis.’ *CMAJ Open* [online] 2 (4), E306-17. available from <http://www.ncbi.nlm.nih.gov/pubmed/25485258> [23 April 2018]

Phelan, S.M., Burgess, D.J., Yeazel, M.W., Hellerstedt, W.L., Griffin, J.M., and van Ryn, M. (2015) ‘Impact of Weight Bias and Stigma on Quality of Care and Outcomes for Patients with Obesity.’ *Obesity Reviews : An Official Journal of the International Association for the Study of Obesity* [online] 16 (4), 319–26. available from <http://www.ncbi.nlm.nih.gov/pubmed/25752756> [25 April 2018]

Prospective Studies Collaboration, Whitlock, G., Lewington, S., Sherliker, P., Clarke, R., Emberson, J., Halsey, J., Qizilbash, N., Collins, R., and Peto, R. (2009) ‘Body-Mass Index and Cause-Specific Mortality in 900 000 Adults: Collaborative Analyses of 57 Prospective Studies’. *The Lancet* [online] 373 (9669), 1083–1096. available from <http://www.ncbi.nlm.nih.gov/pubmed/19299006> [25 April 2018]

Schünemann, H.J., Schünemann, A.H.J., Oxman, A.D., Brozek, J., Glasziou, P., Jaeschke, R., Vist, G.E., Williams, J.W., Kunz, R., Craig, J., Montori, V.M., Bossuyt, P., Guyatt,

G.H., and GRADE Working Group (2008) ‘Grading Quality of Evidence and Strength of Recommendations for Diagnostic Tests and Strategies.’ *BMJ (Clinical Research Ed.)* [online] 336 (7653), 1106–10. available from <http://www.ncbi.nlm.nih.gov/pubmed/18483053> [17 April 2018]

Sterne, J.A.C., Sutton, A.J., Ioannidis, J.P.A., Terrin, N., Jones, D.R., Lau, J., Carpenter, J., Rücker, G., Harbord, R.M., Schmid, C.H., Tetzlaff, J., Deeks, J.J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D.G., Moher, D., and Higgins, J.P.T. (2011) ‘Recommendations for Examining and Interpreting Funnel Plot Asymmetry in Meta-Analyses of Randomised Controlled Trials.’ *BMJ (Clinical Research Ed.)* [online] 343, d4002. available from <http://www.ncbi.nlm.nih.gov/pubmed/21784880> [2 May 2018]

Stevens, J., Truesdale, K.P., McClain, J.E., and Cai, J. (2006) ‘The Definition of Weight Maintenance’. *International Journal of Obesity* [online] 30 (3), 391–399. available from <http://www.nature.com/articles/0803175> [25 April 2018]

The Government Office for Science (2007) *The Foresight Report. Reducing Obesity: Future Choices* [online] available from <https://www.gov.uk/government/publications/reducing-obesity-future-choices>

Thoits, P.A. (2011) ‘Mechanisms Linking Social Ties and Support to Physical and Mental Health’. *Journal of Health and Social Behavior* [online] 52 (2), 145–161. available from <http://journals.sagepub.com/doi/10.1177/0022146510395592> [25 April 2018]

Thoits, P.A. (1995) ‘Stress, Coping, and Social Support Processes: Where Are We? What Next?’ *Journal of Health and Social Behavior* [online] 35, 53. available from <http://www.jstor.org/stable/2626957?origin=crossref> [25 April 2018]

Tur, J.J., Escudero, A.J., Alos, M.M., Salinas, R., Terés, E., Soriano, J.B., Nicola, G.,

Urgelés, J.R., Pagán, A., Cortes, B., González, X., and Burguera, B. (2013) 'One Year Weight Loss in the TRAMOMTANA Study. A Randomized Controlled Trial'. *Clinical Endocrinology* [online] 79 (6), 791–799. available from <http://www.ncbi.nlm.nih.gov/pubmed/23163735> [7 November 2017]

University of York. NHS Centre for Reviews and Dissemination. (2009) *Systematic Reviews : CRD's Guidance for Undertaking Reviews in Health Care*. CRD, University of York

Verheijden, M.W., Bakx, J.C., van Weel, C., Koelen, M.A., and van Staveren, W.A. (2005) 'Role of Social Support in Lifestyle-Focused Weight Management Interventions'. *European Journal of Clinical Nutrition* [online] 59 (S1), S179–S186. available from <http://www.nature.com/articles/1602194> [25 April 2018]

Ware, J.H. (2003) 'Interpreting Incomplete Data in Studies of Diet and Weight Loss'. *New England Journal of Medicine* [online] 348 (21), 2136–2137. available from <http://www.nejm.org/doi/abs/10.1056/NEJMe030054> [2 May 2018]

Westermann, S., Rief, W., Euteneuer, F., and Kohlmann, S. (2015) 'Social Exclusion and Shame in Obesity'. *Eating Behaviors* [online] 17, 74–76. available from <https://www.sciencedirect.com/science/article/pii/S1471015315000033> [25 April 2018]

Williams, G.C., Grow, V.M., Freedman, Z.R., Ryan, R.M., and Deci, E.L. (1996) 'Motivational Predictors of Weight Loss and Weight-Loss Maintenance.' *Journal of Personality and Social Psychology* [online] 70 (1), 115–26. available from <http://www.ncbi.nlm.nih.gov/pubmed/8558405> [25 April 2018]

World Health Organisation (2014) 'Obesity'. *WHO* [online] available from <http://www.who.int/topics/obesity/en/> [25 April 2018]

World Health Organisation (2011) *Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus*. [online] available from <http://www.who.int/diabetes/publications/report-hba1c_2011.pdf> [18 April 2018]

World Health Organisation (2002) *The World Health Report 2002*. [online] available from <http://www.who.int/whr/2002/en/whr02_en.pdf?ua=1> [18 April 2018]

8.0 Appendices

- Appendix 1 – Search Strategies

- Appendix 2 – Data Extraction Form

- Appendix 3 – Risk of Bias Tools

- Appendix 4 – Formula for Combining Groups

- Appendix 5 – Cost-effectiveness Calculations

- Appendix 6 – Ethics Application

Appendix 1

Database	Search		
Wiley Cochrane Library – CENTRAL 2017 Issue 1 Searched 09/02/17	#1	MeSH descriptor: [Obesity] this term only	(8689)
	#2	MeSH descriptor: [Obesity, Abdominal] explode all trees	(140)
	#3	MeSH descriptor: [Obesity, Metabolically Benign] explode all trees	(1)
	#4	MeSH descriptor: [Obesity, Morbid] explode all trees	(833)
	#5	#1 or #2 or #3 or #4	(9583)
	#6	MeSH descriptor: [Life Style] explode all trees	(3896)
	#7	MeSH descriptor: [Diet] explode all trees	(15016)
	#8	MeSH descriptor: [Physical Fitness] explode all trees	(2588)
	#9	MeSH descriptor: [Exercise] explode all trees	(18173)
	#10	MeSH descriptor: [Cognitive Therapy] explode all trees	(7006)
	#11	#6 or #7 or #8 or #9 or #10	(40890)
	#12	MeSH descriptor: [Body Weight Changes] explode all trees	(6385)
	#13	MeSH descriptor: [Body Weight Maintenance] explode all trees	(10)
	#14	MeSH descriptor: [Body Mass Index] explode all trees	(8178)
	#15	#12 or #13 or #14	(12770)
	#16	#5 and #11 and #15	(2736)
	#17	adult not child	(392818)
	#18	humans not animals	(558449)
	#19	#17 and #18	(296030)
	#20	#16 and #19	(1733)
	#21	#20 in Trials	(1677)
EMBSCO CINAHL Searched 09/02/17	#1	exp OBESITY/	42768
	#2	(overweight).ti,ab	10962
	#3	(1 OR 2)	46239
	#4	exp LIFE STYLE/	119156
	#5	exp DIET/	57130
	#6	exp PHYSICAL FITNESS/	9080
	#7	exp EXERCISE/	59050
	#8	exp COGNITIVE THERAPY/	9617
	#9	(4 OR 5 OR 6 OR 7 OR 8)	229099
	#10	exp BODY WEIGHT CHANGES/	54248
	#11	exp BODY MASS INDEX/	38680
	#12	("weight maintenance").ti,ab	364
	#13	(10 OR 11 OR 12)	78769
	#14	"RANDOMIZED CONTROLLED TRIALS"/	30173
	#15	"CLINICAL TRIALS"/	87589
	#16	(random*).ab	119642
	#17	(trial).ti	41913
	#18	(14 OR 15 OR 16 OR 17)	204479
	#19	exp HUMAN/ NOT ANIMAL/	992311
	#20	exp ADULT/ NOT CHILD/	773211
	#21	(19 AND 20)	513829
	#22	(3 AND 9 AND 13 AND 18 AND 21)	1147

EMBSCO EMBASE Searched 09/02/17	#1	OBESITY/ OR "ABDOMINAL OBESITY"/ OR "METABOLICALLY BENIGN OBESITY"/ OR "MORBID OBESITY"/	383282
	#2	(overweight).ti,ab	75540
	#3	(1 OR 2)	390930
	#4	exp "LIFESTYLE MODIFICATION"/	30298
	#5	exp DIET/	329841
	#6	"FITNESS,PHYSICAL"/	45111
	#7	exp EXERCISE/	308790
	#8	exp "COGNITIVE THERAPY"/	43717
	#9	(4 OR 5 OR 6 OR 7 OR 8)	704675
	#10	exp "WEIGHT CHANGE"/	11898
	#11	exp "BODY MASS"/	297611
	#12	("weight maintenance").ti,ab	2409
	#13	(10 OR 11 OR 12)	308003
	#14	'crossover procedure' OR 'double-blind procedure' OR 'randomized controlled trial' OR 'single-blind procedure' OR (random* OR factorial* OR crossover* OR "cross over*" OR "doubl* blind*" OR "singl* blind*" OR assign* OR allocat* OR volunteer*).ab,ti	1849113
	#15	exp HUMAN/ NOT ANIMAL/	18090791
	#16	exp ADULT/ NOT CHILD/	6061223
	#17	(17 AND 18)	5840445
	#18	(3 AND 9 AND 13 AND 14 AND 17)	3425
	#19	18	1107
EMBSCO MEDLINE Searched 09/02/17	#1	OBESITY, ABDOMINAL/ OR OBESITY, METABOLICALLY BENIGN/ OR OBESITY, MORBID/	16660
	#2	OVERWEIGHT/	17160
	#3	(1 OR 2)	33550
	#4	exp LIFE STYLE/	74237
	#5	exp DIET/	231009
	#6	exp PHYSICAL FITNESS/	24317
	#7	exp EXERCISE/	167872
	#8	exp COGNITIVE THERAPY/	20653
	#9	(4 OR 5 OR 6 OR 7 OR 8)	480543

	#10	exp "BODY WEIGHT CHANGES"/	58275
	#11	exp "BODY WEIGHT MAINTENANCE"/	57
	#12	exp "BODY MASS INDEX"/	98685
	#13	(10 OR 11 OR 12)	146738
	#14	(randomized controlled trial).pt	427844
	#15	(controlled clinical trial).pt	316108
	#16	(randomi?ed).ab	444096
	#17	(randomly).ab	264335
	#18	(trial).ti	163498
	#19	"CLINICAL TRIALS AS TOPIC"/	178387
	#20	(14 OR 15 OR 16 OR 17 OR 18 OR 19)	1064759
	#21	exp HUMANS/ NOT ANIMALS/	14549675
	#22	exp ADULT/ NOT CHILDREN/	6162304
	#23	(21 AND 22)	6019184
	#24	(3 AND 9 AND 13 AND 20 AND 23)	1006
ISRCTN register Searched 09/02/17	#1	(obesity OR overweight or “weight loss”)	919

Appendix 2

Population and Setting

	Description *Include comparative information for each group (i.e. intervention and controls) if available*	Location in text *(pg & ¶/fig/table)*
Population description *(from which study participants are drawn)*		
Setting *(including location and social context)*		
Inclusion criteria		
Exclusion criteria		
Method/s of recruitment of participants		
Notes:		

Methods

	Descriptions as stated in report/paper	Location in text *(pg & ¶/fig/table)*
Aim of study		
Unit of allocation *(by individuals, cluster/ groups or body parts)*		

Start date		
End date		
Duration of participation *(from recruitment to last follow-up)*		
Notes:		

Participants

Provide overall data and, if available, comparative data for each intervention or comparison group.

	Description as stated in report/paper	Location in text (pg & ¶/fig/table)
Total no. randomised *(or total pop. at start of study for NRCTs)*		
Clusters *(if applicable, no., type, no. people per cluster)*		
Baseline imbalances		
Withdrawals and exclusions *(if not provided below by outcome)*		

	Description as stated in report/paper	Location in text (pg & ¶/fig/table)
Age		
Gender		
Race/Ethnicity		
BMI Category		
Co-morbidities		
Other relevant sociodemographics		
Subgroups measured		
Notes:		

Intervention groups

Copy and paste table for each intervention and comparison group

Intervention Group 1

	Description as stated in report/paper	Location in text *(pg & ¶/fig/table)*
Group name		
Gender integration or segregation		
Form of lifestyle intervention i.e. diet +/- physical activity		

	Description as stated in report/paper	Location in text *(pg & ¶/fig/table)*
No. randomised to group *(specify whether no. people or clusters)*		
Description *(include sufficient detail for replication, e.g. content, dose, components; if it is a natural experiment, describe the pre-intervention)*		
Duration of treatment period		
Timing *(e.g. frequency, duration of each episode)*		
Delivery *(e.g. mechanism, medium, intensity, fidelity)*		
Interventionist *(e.g. no., profession, training, ethnicity etc. if relevant)*		

	Description as stated in report/paper	Location in text *(pg & ¶/fig/table)*
Co-interventions		
Economic variables *(i.e. intervention cost, changes in other costs as result of intervention)*		
Resource requirements to replicate intervention *(e.g. staff numbers, cold chain, equipment)*		
Notes:		

Outcomes

Copy and paste table for each outcome.

Outcome 1

	Description as stated in report/paper	Location in text *(pg & ¶/fig/table)*
Outcome name		
Time points measured *(specify whether from start or end of intervention)*		
Time points reported		

	Description as stated in report/paper		Location in text *(pg & ¶/fig/table)*
Outcome definition *(with diagnostic criteria if relevant and note whether the outcome is desirable or undesirable if this is not obvious)*			
Person measuring/ reporting			
Unit of measurement *(if relevant)*			
Scales: upper and lower limits *(indicate whether high or low score is good)*			
Is outcome/tool validated?	... *Yes/No/Unclear*		
Imputation of missing data *(e.g. assumptions made for ITT analysis)*			

	Description as stated in report/paper	Location in text *(pg & ¶/fig/table)*
Assumed risk estimate *(e.g. baseline or population risk noted in Background)*		
Notes:		

Results

Copy and paste the appropriate table for each outcome, including additional tables for each time point and subgroup as required.

For randomised or non-randomised trial - Dichotomous outcome

	Description as stated in report/paper				Location in text *(pg & ¶/fig/table)*
Comparison					
Outcome					
Subgroup					
Time point *(specify whether from start or end of intervention)*					
Results *Note whether:* ... *post-intervention OR*	Intervention		Comparison		
	No. events	No. participants	No. events	No. participants	

	Description as stated in report/paper				Location in text *(pg & ¶/fig/table)*
... *change from baseline* *And whether* ... *Adjusted OR* ...*Unadjusted*					
Baseline data	Intervention		Comparison		
	No. events	No. participants	No. events	No. participants	
No. missing participants and reasons					
No. participants moved from other group and reasons					
Any other results reported					
Unit of analysis *(e.g. by individuals, health professional, practice, hospital, community)*					

	Description as stated in report/paper		Location in text *(pg & ¶/fig/table)*
Statistical methods used and appropriateness of these methods *(e.g. adjustment for correlation)*			
Reanalysis required? *(if yes, specify why, e.g. correlation adjustment)*	... *Yes/No/Unclear*		
Reanalysis possible?	... *Yes/No/Unclear*		
Reanalysed results			
Notes:			

For randomised or non-randomised trial - Continuous outcome

	Description as stated in report/paper	Location in text *(pg & ¶/fig/table)*
Comparison		
Outcome		
Subgroup		
Time point *(specify whether from start or end of intervention)*		

	Description as stated in report/paper	Location in text *(pg & ¶/fig/table)*
Post-intervention or change from baseline?		

Results *Note whether:* ... *post-intervention* *OR* ... *change from baseline* *And whether* ... *Adjusted* *OR* ...*Unadjusted*	Intervention			Comparison			
	Mean	SD (or other variance)	No. participants	Mean	SD (or other variance)	No. participants	

Baseline data	Intervention			Comparison			
	Mean	SD (or other variance)	No. participants	Mean	SD (or other variance)	No. participants	

No. missing participants and reasons			

	Description as stated in report/paper		Location in text *(pg & ¶/fig/table)*
No. participants moved from other group and reasons			
Any other results reported			
Unit of analysis *(e.g. by individuals, health professional, practice, hospital, community)*			
Statistical methods used and appropriateness of these methods *(e.g. adjustment for correlation)*			
Reanalysis required? *(if yes, specify why)*	... *Yes/No/Unclear*		
Reanalysis possible?	... *Yes/No/Unclear*		
Reanalysed results			
Notes:			

For randomised or non-randomised trial - Other outcome

	Description as stated in report/paper				Location in text *(pg & ¶/fig/table)*
Comparison					
Outcome					
Subgroup					
Time point *(specify whether from start or end of intervention)*					
Type of outcome					
Results	Intervention result	SD (or other variance)	Control result	SD (or other variance)	
	Overall results		SE (or other variance)		
No. participant	Intervention		Control		
No. missing participants and reasons					
No. participants moved from other group and reasons					
Any other results reported					

	Description as stated in report/paper		Location in text *(pg & ¶/fig/table)*
Unit of analysis *(e.g. by individuals, health professional, practice, hospital, community)*			
Statistical methods used and appropriateness of these methods			
Reanalysis required? *(if yes, specify why)*	...		
Reanalysis possible?	...		
Reanalysed results			
Notes:			

Applicability

Have important populations been excluded from the study? *(consider disadvantaged populations, and possible differences in the intervention effect)*	... *Yes/No/Unclear*	

Is the intervention likely to be aimed at disadvantaged groups? *(e.g. lower socioeconomic groups)*	... *Yes/No/Unclear*	
Does the study directly address the review question? *(any issues of partial or indirect applicability)*	... *Yes/No/Unclear*	
Notes:		

Other information

	Description as stated in report/paper	Location in text *(pg & ¶/fig/table)*
Key conclusions of study authors		
Notes:		

Appendix 3

Short Title	Attrition Bias	Selection Bias	Detection Bias	Performance Bias	Reporting Bias
Appel (2011)	**Incomplete outcome data*** Low risk [Info] - *ITT analysis - missing data imputed - all participants accounted for in CONSORT flow diagram - attrition balanced across groups - although a 'per-protocol' analysis was used before a protocol-defined censoring event (pregnancy, bariatric surgery or amputation), data prior to this event was still included. This is low risk as this was pre-specified in the protocol and weight loss outcome/patient-reported outcomes would be influenced by these objectively assessed physiological*	**Random sequence generation** Low risk [Info] - *web-based randomisation - block randomisation* **Allocation concealment** Low risk [Info] - *web based programme (central allocation)*	**Blinding of outcome assessment*** Low risk [Info] - *outcome assessors were blinded to allocation*	**Blinding of participants and personnel*** Unclear [Info] *Due to the nature of randomisation into remote vs in-person support, it is not clear if the participants would be aware of what arm they had been randomised to. Group intervention provided by John Hopkins University staff, whereas Healthways provided remote-support - different staff groups providing intervention, but un-clear if they knew which was the more intensive intervention.*	**Selective reporting** High risk [Info] - *Dietary outcomes not reported i.e. fruit and vegetable. fat intake - Binge eating score not reported - Social support not reported The extensive outcomes listed in the protocol versus the outcomes reported suggests that the authors were using a 'wide net' and reporting significant secondary outcomes only*

	changes				
Ash (2006)	**Incomplete outcome data*** High risk [Info] - *attrition was poor and imbalanced - 47% group vs 68% individual included in analysis - although authors stated that data was handled as ITT, this analysis was not reported - no imputation for missing data*	**Random sequence generation** Low risk [Info] - *random table allocation* **Allocation concealment** Unclear [Info] - *project manager carried out both recruitment and randomisation of participants - may have led to selection bias if this was through direct contact*	**Blinding of outcome assessment*** High risk [Info] - *dietitians (caregivers) are collecting the outcome data*	**Blinding of participants and personnel*** High risk [Info] - *participants and personnel aware of allocation*	**Selective reporting** High risk [Info] - *protocol not available - in the paper authors stated data was collected at 4 time points, however only presented 3 time points - did not present data, or significance of data, between baseline and follow-up for SWLS or GHQ-12 - only presented baseline data*
Heshka (2003)	**Incomplete outcome data*** Low risk [Info] - *ITT analysis - missing data imputed - all participants accounted for in CONSORT flow diagram - attrition balanced*	**Random sequence generation** Low risk [Info] - *random number table* **Allocation concealment** Unclear [Info] -	**Blinding of outcome assessment*** Unclear [Info] - *no detail about who collected the outcome data*	**Blinding of participants and personnel*** High risk [Info] - *participants and personnel were aware of allocation*	**Selective reporting** Unclear [Info] - *study protocol not available*

	across groups	*central allocation - however, insufficient detail about concealment of envelopes - screening and test results were reviewed by the central allocation team prior to randomisation*			
Tur (2013)	**Incomplete outcome data*** High risk [Info] - *no imputations for missing data - 68% group vs 38% individual*	**Random sequence generation** Unclear [Info] - *insufficient detail of "computer derived" randomisation* **Allocation concealment** Unclear [Info] - *insufficient detail of randomisation process "student was not blinded"*	**Blinding of outcome assessment*** High risk [Info] - *outcome assessors were not blinded*	**Blinding of participants and personnel*** High risk [Info] - *study was not blinded to personnel or participants*	**Selective reporting** High risk [Info] - *protocol published - outcomes of anxiety, depression and SF-36 were not reported but were pre-specified as secondary outcomes in the protocol*
Jolly (2011)	**Incomplete outcome data***	**Random sequence**	**Blinding of outcome**	**Blinding of participants**	**Selective reporting**

	Low risk [Info] - *ITT analysis - missing data imputed - all participants account for in CONSORT flow diagram - overall, comparable attrition across groups*	**generation** Low risk [Info] - *independent statistician prepared the randomisation sequences* **Allocation concealment** Low risk [Info] - *a central call centre system was used - call centre staff were blinded to the sequence by using opaque envelopes in consecutively numbered order*	**assessment*** High risk [Info] - *primary outcome not blinded to personnel or participant - secondary outcome at 12 months was blinded to the assessor, but not participant - IPAQ is not suitable to detect differences in physical activity over time (only useful for cross-sectional stuies)*	**and personnel*** High risk [Info] - *participants and personnel were aware of allocation*	Low risk [Info] - *protocol available (Jolly et al., 2010) and all pre-specified outcome data has been reported in the manuscript*
Long (1983)	**Incomplete outcome data*** High risk [Info] - *no imputation for missing data - due to small sample size, unbalanced attrition between individual vs group interventions (> missing data for individual*	**Random sequence generation** Unclear [Info] - *insufficient information about randomisation process* **Allocation concealment** Unclear [Info] - *this is*	**Blinding of outcome assessment*** High risk [Info] – *interventionists were the outcome assessors*	**Blinding of participants and personnel*** High risk [Info] - *participants and personnel were aware of allocation*	**Selective reporting** Unclear [Info] - *no protocol available*

	intervention)	*not described*			
McRobbie (2016)	**Incomplete outcome data*** Low risk [Info] - *ITT analysis - missing data was imputed - 88% vs 90% included in primary analysis - all participants accounted for in consort flow diagram*	**Random sequence generation** Low risk [Info] - *web-based application for randomisation (produced by statistician)* **Allocation concealment** Low risk [Info] - *central allocation*	**Blinding of outcome assessment*** Low risk [Info] - *outcome assessors were blinded*	**Blinding of participants and personnel*** High risk [Info] - *participants and personnel aware of allocation*	**Selective reporting** Low risk [Info] - *protocol available - transparency about protocol ammendments : "primary and secondary outcomes were clarified" on 17/05/12. Recruitment commenced Sept 2012, therefore reported outcomes weren't altered based upon the available data*
Jebb (2011)	**Incomplete outcome data*** Low risk [Info] - *ITT analysis - missing data imputed - all participants accounted for in CONSORT flowchart - 61% group vs. 54% individual*	**Random sequence generation** Low risk [Info] - *computer generated randomisation sequence* **Allocation concealment**	**Blinding of outcome assessment*** High risk [Info] - *outcome assessment was not masked*	**Blinding of participants and personnel*** High risk [Info] - *personnel and participants were not masked*	**Selective reporting** Unclear [Info] - *protocol no longer available online*

	data completion, which is fairly equivalent	Low risk [Info] - *allocation concealed through central randomisation*			

Appendix 4

mean1		**mean2**		*mean12*	*=((C60*C61)+(E60*E61))/I61*
n1		**n2**		*n12*	*=C61+E61*
sd1		**sd2**		*sd12*	*=SQRT(I63)*
var1	=C62^2	var2	=E62^2	var12	=(I64-(I61*(I60^2)))/(I61-1)
sumsq1	=(C61-1)*(C63+((C61/(C61-1))*(C60^2)))	sumsq2	=(E61-1)*(E63+((E61/(E61-1))*(E60^2)))	sumsq12	=C64+E64

Appendix 5

Formula:

Cost per QALY = [a] multiplied by [b]

Study	Country	n=	Effectiveness		Cost		ICER (£) 2016
			Mean Difference (QALY) [a]	p-value	Mean Difference (£) (year) [b]	p-value	
Jebb	Australia	243	0.021	0.006	228 (2011)	-	12,143
	Germany	238	0.009	0.222	116 (2011)	-	14,416
	UK	178	0.015	0.033	109 (2011)	-	8,128
McRobbie	UK	179	0.0104	0.088	81 (2012)	0.787	8,439

Converted into estimated cost for 2016 according to Bank of England inflation rate using the Bank of England online 'Inflation Calculator'. Available from: https://www.bankofengland.co.uk/monetary-policy/inflation [2 February 2018]

Appendix 6

Low Risk Research Ethics Approval

Project Title

The Influence of Gender on the Efficacy of Group and Individual Lifestyle Interventions for Overweight and Obese Adults: a Systematic Review and Meta-Analysis of Randomised Controlled Trials

Record of Approval

Principal Investigator

I **request an ethics peer review** and confirm that I have answered all relevant questions in this checklist honestly.	X
I confirm that I will carry out the project in the ways described in this checklist. I will immediately suspend research and request new ethical approval if the project subsequently changes the information I have given in this checklist.	X
I confirm that I, and all members of my research team (if any), have read and agreed to abide by the Code of Research Ethics issued by the relevant national learned society.	X
I confirm that I, and all members of my research team (if any), have read and agreed to abide by the University's Research Ethics, Governance and Integrity Framework.	X

Name: Sally Abbott .....

Date: 27/01/2017 .....

Student's Supervisor (if applicable)

I have read this checklist and confirm that it covers all the ethical issues raised by this project fully and frankly. I also confirm that these issues have been discussed with the student and will continue to be reviewed in the course of supervision.

Name: Deborah Lycett .....

Date: 31/01/2017 .....

Reviewer (if applicable)

Date of approval by anonymous reviewer: 31/01/2017

Low Risk Research Ethics Approval Checklist

Project Information

Project Ref	P51006
Full name	Sally Abbott
Faculty	Faculty of Health and Life Sciences
Department	FRC Technology Enabled Health Research (CTEHR)
Supervisor	Deborah Lycett
Module Code	M001RDC
EFAAF Number	
Project title	The Influence of Gender on the Efficacy of Group and Individual Lifestyle Interventions for Overweight and Obese Adults: a Systematic Review and Meta-Analysis of Randomised Controlled Trials
Date(s)	03/02/2017 - 04/05/2018
Created	27/01/2017 14:00

Project Summary

Research questions:

(1)How effective are group compared to individual lifestyle interventions at improving obesity-related clinical outcomes?

(2)Does gender influence the efficacy of group compared to individual lifestyle interventions at improving obesity-related clinical outcomes?

(3)Is there a difference in the efficacy of gender-integrated and gender-segregated group lifestyle interventions compared to individual lifestyle interventions at improving obesity-related clinical outcomes?

Objectives:

1.To identify publications that compare group versus individual lifestyle interventions for adult obesity

2.To critically appraise the methodological quality of identified publications

3.To synthesise data from identified publications in order:

To establish whether there is a difference in the efficacy of group versus individual lifestyle interventions

To establish whether there is a relationship between gender and the efficacy of group versus individual lifestyle interventions

To establish if there is a relationship between gender integration and the efficacy of group lifestyle interventions versus individual lifestyle interventions

4.Identify priorities for future

Names of Co-Investigators and their organisational affiliation (place of study/employer)	
Is the project self-funded?	YES
Who is funding the project?	NIHR
Has the funding been confirmed?	YES
Are you required to use a Professional Code of Ethical Practice appropriate to your discipline?	NO
Have you read the Code?	NO

Project Details

What is the purpose of the project?	the aim of this research is to not only conduct an updated systematic review of the evidence of the efficacy of group and individual lifestyle interventions; but to also examine in detail the influence of gender on intervention effectiveness.
What are the planned or desired outcomes?	The findings of this systematic review will inform future research into the role of gender-specific interventions in addressing obesity and could provide new knowledge to inform clinical guidelines and clinical practice in the management of adult obesity
Explain your research design	A systematic review and meta-analysis.
Outline the principal methods you will use	Meta-analysis is planned as long as a minimum of 2 RCTs have been identified. A visual test for heterogeneity will be used to assess the overlap in confidence intervals for each effect estimate on a forest plot. If the overlap is poor or there are outliers, statistical heterogeneity may be likely. A test for statistical heterogeneity will then be performed using I2 statistics. If statistical heterogeneity exists, as defined by I2]50%, this will be addressed by conducting a random-effects model in meta-analysis followed by a further sub-group analysis of homogenous studies. Substantial heterogeneity would also mean that a narrative synthesis would take place. If a narrative approach is taken, information will be presented in the text and tables to summarise and explain the characteristics and findings of the included studies. The narrative synthesis will explore the relationship and findings both within and between the included studies.
Are you proposing to use an external research instrument, validated scale or follow a published research method?	NO
If yes, please give details of what you are using	
Will your research involve consulting individuals who support, or literature, websites or similar material which advocates, any of the following: terrorism, armed struggles, or political, religious or other forms of activism considered illegal under UK law?	NO

Are you dealing with Secondary Data? (e.g. sourcing info from websites, historical documents)	YES
Are you dealing with Primary Data involving people? (e.g. interviews, questionnaires, observations)	NO
Are you dealing with personal or sensitive data?	NO
Is the project solely desk based? (e.g. involving no laboratory, workshop or off-campus work or other activities which pose significant risks to researchers or participants)	YES
Are there any other ethical issues or risks of harm raised by the study that have not been covered by previous questions?	NO
If yes, please give further details	

External Ethical Review

Question		Yes	No
1	Will this study be submitted for ethical review to an external organisation? (e.g. Another University, Social Care, National Health Service, Ministry of Defence, Police Service and Probation Office)		X
	If YES, name of external organisation		
2	Will this study be reviewed using the IRAS system?		X
3	Has this study previously been reviewed by an external organisation?		X

Risk of harm, potential harm and disclosure of harm

Question			Yes	No
1	Is there any significant risk that the study may lead to physical harm to participants or researchers?			X
	If YES, please explain how you will take steps to reduce or address those risks			
2	Is there any significant risk that the study may lead to psychological or emotional distress to participants?			X
	If YES, please explain how you will take steps to reduce or address those risks			
3	Is there any risk that the study may lead to psychological or emotional distress to researchers?			X
	If YES, please explain how you will take steps to reduce or address those risks			
4	Is there any risk that your study may lead or result in harm to the reputation of participants, researchers, or their employees, or any associated persons or organisations?			X
	If YES, please explain how you will take steps to reduce or address those risks			
5	Is there a risk that the study will lead to participants to disclose evidence of previous criminal offences, or their intention to commit criminal offences?			X
	If YES, please explain how you will take steps to reduce or address those risks			
6	Is there a risk that the study will lead participants to disclose evidence that children or vulnerable adults are being harmed, or at risk or harm?			X
	If YES, please explain how you will take steps to reduce or address those risks			
7	Is there a risk that the study will lead participants to disclose evidence of serious risk of other types of harm?			X
	If YES, please explain how you will take steps to reduce or address those risks			
8	Are you aware of the CU Disclosure protocol?		X	

Online and Internet Research

Question			Yes	No
1	Will any part of your study involve collecting data by means of electronic media (e.g. the Internet, e-mail, Facebook, Twitter, online forums, etc)?		X	
	If YES, please explain how you will obtain permission to collect data by this means	Using EBSCO databases - Coventry University has a license		
2	Is there a possibility that the study will encourage children under 18 to access inappropriate websites, or correspond with people who pose risk of harm?			X
	If YES, please explain further			
3	Will the study incur any other risks that arise specifically from the use of electronic media?			X
	If YES, please explain further			
4	Will you be using survey collection software (e.g. BoS, Filemaker)?			X
	If YES, please explain which software			
5	Have you taken necessary precautions for secure data management, in accordance with data protection and CU Policy?		X	
	If NO	please explain why not		
	If YES	Specify location where data will be stored	EPPI-Reviewer 4 is a web based, secure data storage system	
		Planned disposal date	01/01/2020	
		If the research is funded by an external organisation, are there any requirements for storage and disposal?		X
		If YES, please specify details		